



Europäisches Patentamt  
European Patent Office  
Office européen des brevets



(11)

**EP 1 355 149 A2**

(12)

**EUROPEAN PATENT APPLICATION**

(43) Date of publication:  
22.10.2003 Bulletin 2003/43

(51) Int Cl.7: **G01N 33/48**, C07H 21/04,  
C12Q 1/68

(21) Application number: **03252022.3**

(22) Date of filing: **31.03.2003**

(84) Designated Contracting States:  
**AT BE BG CH CY CZ DE DK EE ES FI FR GB GR**  
**HU IE IT LI LU MC NL PT RO SE SI SK TR**  
Designated Extension States:  
**AL LT LV MK**

(30) Priority: **29.03.2002 US 368798 P**

(71) Applicant: **Ortho-Clinical Diagnostics, Inc.**  
**Rochester, NY 14626-5101 (US)**

(72) Inventor: **Wang, Yixin**  
**San Diego, CA 92130 (US)**

(74) Representative: **Mercer, Christopher Paul et al**  
**Carpmaels & Ransford**  
**43, Bloomsbury Square**  
**London WC1A 2RA (GB)**

(54) **Assessing colorectal cancer**

(57) A method of assessing the presence or absence of colorectal cancer or the likely condition of a person believed to have colorectal cancer is conducted by analyzing the expression of a group of genes. Gene

expression profiles in a variety of medium such as microarrays are included as are kits that contain them.

**EP 1 355 149 A2**

**Description****BACKGROUND**

- 5 [0001] This application claims the benefit of U.S. Provisional Application No. 60/368,798 filed on March 29, 2002.
- [0002] This invention relates to diagnostics and prognostics for colorectal cancer based on the gene expression profiles of biological samples.
- 10 [0003] Colorectal cancer is a heterogeneous disease, consisting of tumors thought to emerge through three major molecular mechanisms: 1) mutations in the adenomatous polyposis coli (APC) gene, or the  $\beta$ -catenin gene, combined with chromosomal instability, 2) mutations in DNA mismatch repair genes, such as MLH1, MSH2, PMS1, PMS2 and MSH6, associated with microsatellite instability and mutations in genes containing short repeats, and 3) gene silencing induced by hypermethylation of the promoter regions of tumor suppressor genes. The genetic complement of individual colorectal cancers is likely to include different combinations of genetic instability, specific mutations, and gene silencing. Chromosomal instability (CIN) is a common feature of cancers in general. It implies an aneuploid phenotype, in which whole chromosomes or large parts of them are being lost or gained. Microsomal instability (MIN) is found in diploid tumors with an increased mutation rate in short repeats. Both forms of genetic instability are common in colorectal cancer.
- 15 [0004] Colorectal cancers thus have complex origins and involve a number of interactions in different biological pathways. Serum markers, histological, and cytological examinations historically used to assist in providing diagnostic, prognostic, or therapy monitoring decisions often do not have desired reliability. Likewise, while use of a single genetic marker (e.g., increased expression of a particular gene) may be beneficial, the diversity of the cancers make it more likely that a portfolio of genetic markers is the best approach.
- 20

**SUMMARY OF THE INVENTION**

- 25 [0005] The invention is a method of assessing the presence or absence of colorectal cancer or the likely condition of a person believed to have colorectal cancer. In the method, a gene expression profile of a patient sample is analyzed to determine whether a patient has a colorectal cancer, whether a patient does not have colorectal cancer, whether a patient is likely to get colorectal cancer, or the response to treatment of a patient being treated for colorectal cancer.
- 30 [0006] Articles used in practicing the methods are also an aspect of the invention. Such articles include gene expression profiles or representations of them that are fixed in machine-readable media such as computer readable media.
- [0007] Articles used to identify gene expression profiles can also include substrates or surfaces, such as microarrays, to capture and/or indicate the presence, absence, or degree of gene expression.
- 35

**DETAILED DESCRIPTION**

- [0008] The mere presence or absence of particular nucleic acid sequences in a tissue sample has only rarely been found to have diagnostic or prognostic value. Information about the expression of various proteins, peptides or mRNA, on the other hand, is increasingly viewed as important. The mere presence of nucleic acid sequences having the potential to express proteins, peptides, or mRNA (such sequences referred to as "genes") within the genome by itself is not determinative of whether a protein, peptide, or mRNA is expressed in a given cell. Whether or not a given gene capable of expressing proteins, peptides, or mRNA does so and to what extent such expression occurs, if at all, is determined by a variety of complex factors. Irrespective of difficulties in understanding and assessing these factors, assaying gene expression can provide useful information about the occurrence of important events such as tumorigenesis, metastasis, apoptosis, and other clinically relevant phenomena. Relative indications of the degree to which genes are active or inactive can be found in gene expression profiles. The gene expression profiles of this invention are used to diagnose and treat patients for colorectal cancer.
- 40
- 45 [0009] Sample preparation requires the collection of patient samples. Patient samples used in the inventive method are those that are suspected of containing diseased cells such as epithelial cells taken from a colon sample or from surgical margins. One useful technique for obtaining suspect samples is Laser Capture Microdissection (LCM). LCM technology provides a way to select the cells to be studied, minimizing variability caused by cell type heterogeneity. Consequently, moderate or small changes in gene expression between normal and cancerous cells can be readily detected. In a preferred method, the samples comprise circulating epithelial cells extracted from peripheral blood.
- 50
- 55 These can be obtained according to a number of methods but the most preferred method is the magnetic separation technique described in U.S. Patent 6,136,182 assigned to Immunivest Corp which is incorporated herein by reference. Once the sample containing the cells of interest has been obtained, RNA is extracted and amplified and a gene expression profile is obtained, preferably via micro-array, for genes in the appropriate portfolios.

- [0010]** Preferred methods for establishing gene expression profiles include determining the amount of RNA that is produced by a gene that can code for a protein or peptide. This is accomplished by reverse transcriptase PCR (RT-PCR), competitive RT-PCR, real time RT-PCR, differential display RT-PCR, Northern Blot analysis and other related tests. While it is possible to conduct these techniques using individual PCR reactions, it is best to amplify complementary DNA (cDNA) or complementary RNA (cRNA) produced from mRNA and analyze it via microarray. A number of different array configurations and methods for their production are known to those of skill in the art and are described in U.S. Patents such as: 5,445,934; 5,532,128; 5,556,752; 5,242,974; 5,384,261; 5,405,783; 5,412,087; 5,424,186; 5,429,807; 5,436,327; 5,472,672; 5,527,681; 5,529,756; 5,545,531; 5,554,501; 5,561,071; 5,571,639; 5,593,839; 5,599,695; 5,624,711; 5,658,734; and 5,700,637; the disclosures of which are incorporated herein by reference.
- [0011]** Microarray technology allows for the measurement of the steady-state mRNA level of thousands of genes simultaneously thereby presenting a powerful tool for identifying effects such as the onset, arrest, or modulation of uncontrolled cell proliferation. Two microarray technologies are currently in wide use. The first are cDNA arrays and the second are oligonucleotide arrays. Although differences exist in the construction of these chips, essentially all downstream data analysis and output are the same. The product of these analyses are typically measurements of the intensity of the signal received from a labeled probe used to detect a cDNA sequence from the sample that hybridizes to a nucleic acid sequence at a known location on the microarray. Typically, the intensity of the signal is proportional to the quantity of cDNA, and thus mRNA, expressed in the sample cells. A large number of such techniques are available and useful. Preferred methods for determining gene expression can be found in US Patents 6,271,002 to Linsley, et al.; 6,218,122 to Friend, et al.; 6,218,114 to Peck, et al.; and 6,004,755 to Wang, et al., the disclosure of each of which is incorporated herein by reference.
- [0012]** Analysis of the expression levels is conducted by comparing such intensities. This is best done by generating a ratio matrix of the expression intensities of genes in a test sample versus those in a control sample. For instance, the gene expression intensities from a diseased tissue can be compared with the expression intensities generated from normal tissue of the same type (e.g., diseased colon tissue sample vs. normal colon tissue sample). A ratio of these expression intensities indicates the fold-change in gene expression between the test and control samples.
- [0013]** Gene expression profiles can also be displayed in a number of ways. The most common method is to arrange a raw fluorescence intensities or ratio matrix into a graphical dendrogram where columns indicate test samples and rows indicate genes. The data is arranged so genes that have similar expression profiles are proximal to each other. The expression ratio for each gene is visualized as a color. For example, a ratio less than one (indicating down-regulation) may appear in the blue portion of the spectrum while a ratio greater than one (indicating up-regulation) may appear as a color in the red portion of the spectrum. Commercially available computer software programs are available to display such data including "GENESPRING" from Silicon Genetics, Inc. and "DISCOVERY" and "INFER" software from Partek, Inc.
- [0014]** Modulated genes used in the methods of the invention are shown in Table 1. The genes that are differentially expressed are shown as being either up regulated or down regulated in diseased cells. Up regulation and down regulation are relative terms meaning that a detectable difference (beyond the contribution of noise in the system used to measure it) is found in the amount of expression of the genes relative to some baseline. In this case, the baseline is the measured gene expression of a normal cell. The genes of interest in the diseased cells are then either up regulated or down regulated relative to the baseline level using the same measurement method. Diseased, in this context, refers to an alteration of the state of a body that interrupts or disturbs, or has the potential to disturb, proper performance of bodily functions as occurs with the uncontrolled proliferation of cells. Someone is diagnosed with a disease when some aspect of that person's genotype or phenotype is consistent with the presence of the disease. However, the act of conducting a diagnosis or prognosis includes the determination disease/status issues such as therapy monitoring. In therapy monitoring, clinical judgments are made regarding the effect of a given course of therapy by comparing the expression of genes over time to determine whether the gene expression profiles have changed or are changing to patterns more consistent with normal tissue.
- [0015]** Preferably, levels of up and down regulation are distinguished based on fold changes of the intensity measurements of hybridized microarray probes. A 2.0 fold difference is preferred for making such distinctions or a p-value less than .05. That is, before a gene is said to be differentially expressed in diseased versus normal cells, the diseased cell is found to yield at least 2 more, or 2 times less intensity than the normal cells. The greater the fold difference, the more preferred is use of the gene as a diagnostic. Genes selected for the gene expression profiles of the instant invention have expression levels that result in the generation of a signal that is distinguishable from those of the normal or non-modulated genes by an amount that exceeds background using clinical laboratory instrumentation.
- [0016]** Statistical values can be used to confidently distinguish modulated from non-modulated genes and noise. Statistical tests find the genes most significantly different between diverse groups of samples. The Student's t-test is an example of a robust statistical test that can be used to find significant differences between two groups. The lower the p-value, the more compelling the evidence that the gene is showing a difference between the different groups. Nevertheless, since microarrays measure more than one gene at a time, tens of thousands of statistical tests may be

asked at one time. Because of this, there is likelihood to see small p-values just by chance and adjustments for this using a Sidak correction as well as a randomization/permutation experiment can be made. A p-value less than .05 by the t-test is evidence that the gene is significantly different. More compelling evidence is a p-value less than .05 after the Sidak correct is factored in. For a large number of samples in each group, a p-value less than 0.05 after the randomization/permutation test is the most compelling evidence of a significant difference.

**[0017]** Another parameter that can be used to select genes that generate a signal that is greater than that of the non-modulated gene or noise is the use of a measurement of absolute signal difference. Preferably, the signal generated by the modulated gene expression is at least 20% different than those of the normal or non-modulated gene (on an absolute basis). It is even more preferred that such genes produce expression patterns that are at least 30% different than those of normal or non-modulated genes.

**[0018]** Genes can be grouped so that information obtained about the set of genes in the group provides a sound basis for making a clinically relevant judgment such as a diagnosis, prognosis, or treatment choice. These sets of genes make up the portfolios of the invention. In this case, the judgments supported by the portfolios involve colorectal cancer. Portfolios of gene expression profiles can be comprised of combinations of genes described in Example 3. As with most diagnostic markers, it is often desirable to use the fewest number of markers sufficient to make a correct medical judgment. This prevents a delay in treatment pending further analysis as well inappropriate use of time and resources. In this case, such a minimal portfolio can be comprised of a combination of genes from Example 4.

**[0019]** Preferably, portfolios are established such that the combination of genes in the portfolio exhibit improved sensitivity and specificity relative to individual genes or randomly selected combinations of genes. In the context of the instant invention, the sensitivity of the portfolio can be reflected in the fold differences exhibited by a gene's expression in the diseased state relative to the normal state. Specificity can be reflected in statistical measurements of the correlation of the signaling of gene expression with the condition of interest. For example, standard deviation can be used as such a measurement. In considering a group of genes for inclusion in a portfolio, a small standard deviation in expression measurements correlates with greater specificity. Other measurements of variation such as correlation coefficients can also be used in this capacity. The most preferred method of establishing gene expression portfolios is through the use of optimization algorithms such as the mean variance algorithm widely used in establishing stock portfolios. This method is described in detail in the co-pending patent application entitled "Portfolio Selection" by Tim Jatkoe, et. al., of equal date hereto. Essentially, the method calls for the establishment of a set of inputs (stocks in financial applications, expression as measured by intensity here) that will optimize the return (e.g., signal that is generated) one receives for using it while minimizing the variability of the return. Many commercial software programs are available to conduct such operations. "Wagner Associates Mean-Variance Optimization Application", referred to as "Wagner Software" throughout this specification, is preferred. This software uses functions from the "Wagner Associates Mean-Variance Optimization Library" to determine an efficient frontier and optimal portfolios in the Markowitz sense is preferred.

**[0020]** Use of this type of software requires that microarray data be transformed so that it can be treated as an input in the way stock return and risk measurements are used when the software is used for its intended financial analysis purposes. For example, when Wagner Software is employed in conjunction with microarray intensity measurements the following data transformation method is employed.

**[0021]** Genes are first pre-selected by identifying those genes whose expression shows at least some minimal level of differentiation. The preferred pre-selection process is conducted as follows. A baseline class is selected. Typically, this will comprise genes from a population that does not have the condition of interest. For example, if one were interested in selecting a portfolio of genes that are diagnostic for breast cancer, samples from patients without breast cancer can be used to make the baseline class. Once the baseline class is selected, the arithmetic mean and standard deviation is calculated for the indicator of gene expression of each gene for baseline class samples. This indicator is typically the fluorescent intensity of a microarray reading. The statistical data computed is then used to calculate a baseline value of  $(X \times \text{Standard Deviation} + \text{Mean})$  for each gene. This is the baseline reading for the gene from which all other samples will be compared. X is a stringency variable selected by the person formulating the portfolio. Higher values of X are more stringent than lower. Preferably, X is in the range of .5 to 3 with 2 to 3 being more preferred and 3 being most preferred.

**[0022]** Ratios between each experimental sample (those displaying the condition of interest) versus baseline readings are then calculated. The ratios are then transformed to base 10 logarithmic values for ease of data handling by the software. This enables down regulated genes to display negative values necessary for optimization according to the Markman mean-variance algorithm using the Wagner Software.

**[0023]** The preprocessed data comprising these transformed ratios are used as inputs in place of the asset return values that are normally used in the Wagner Software when it is used for financial analysis purposes.

**[0024]** Once an efficient frontier is formulated, an optimized portfolio is selected for a given input level (return) or variance that corresponds to a point on the frontier. These inputs or variances are the predetermined standards set by the person formulating the portfolio. Stated differently, one seeking the optimum portfolio determines an acceptable

input level (indicative of sensitivity) or a given level of variance (indicative of specificity) and selects the genes that lie along the efficient frontier that correspond to that input level or variance. The Wagner Software can select such genes when an input level or variance is selected. It can also assign a weight to each gene in the portfolio as it would for a stock in a stock portfolio.

- 5 [0025] Determining whether a sample has the condition for which the portfolio is diagnostic can be conducted by comparing the expression of the genes in the portfolio for the patient sample with calculated values of differentially expressed genes used to establish the portfolio. Preferably, a portfolio value is first generated by summing the multiples of the intensity value of each gene in the portfolio by the weight assigned to that gene in the portfolio selection process. A boundary value is then calculated by (Y\*standard deviation + mean of the portfolio value for baseline groups) where
- 10 Y is a stringency value having the same meaning as X described above. A sample having a portfolio value greater than the portfolio value of the baseline class is then classified as having the condition. If desired, this process can be conducted iteratively in accordance with well known statistical methods for improving confidence levels.
- [0026] Optionally one can reiterate this process until best prediction accuracy is obtained.
- [0027] The process of portfolio selection and characterization of an unknown is summarized as follows:

- 15
1. Choose baseline class.
  2. Calculate mean, and standard deviation of each gene for baseline class samples.
  3. Calculate (X\*Standard Deviation + Mean) for each gene. This is the baseline reading from which all other samples will be compared. X is a stringency variable with higher values of X being more stringent than lower.
  - 20 4. Calculate ratio between each Experimental sample versus baseline reading calculated in step 3.
  5. Transform ratios such that ratios less than 1 are negative (eg.using Log base 10). (Down regulated genes now correctly have negative values necessary for MV optimization).
  6. These transformed ratios are used as inputs in place of the asset returns that are normally used in the software application.
  - 25 7. The software will plot the efficient frontier and return an optimized portfolio at any point along the efficient frontier.
  8. Choose a desired return or variance on the efficient frontier.
  9. Calculate the Portfolio's Value for each sample by summing the multiples of each gene's intensity value by the weight generated by the portfolio selection algorithm.
  - 30 10. Calculate a boundary value by adding the mean Portfolio Value for Baseline groups to the multiple of Y and the Standard Deviation of the Baseline's Portfolio Values. Values greater than this boundary value shall be classified as the Experimental Class.
  11. Optionally one can reiterate this process until best prediction accuracy is obtained.

- [0028] Alternatively, genes can first be pre-selected by identifying those genes whose expression shows some minimal level of differentiation. The pre-selection in this alternative method is preferably based on a threshold given by
- 35

$$1 \leq \left| \frac{(\mu_t - \mu_n)}{(\sigma_t + \sigma_n)} \right|,$$

- 40 where  $\mu_t$  is the mean of the subset known to possess the disease or condition,  $\mu_n$  is the mean of the subset of normal samples, and  $\sigma_t + \sigma_n$  represent the combined standard deviations. A signal to noise cutoff can also be used by pre-selecting the data according to a relationship such as

45

$$0.5 \leq \left| \frac{(\mu_t - MAX_n)}{(\sigma_t + \sigma_n)} \right|$$

- 50 This ensures that genes that are pre-selected based on their differential modulation are differentiated in a clinically significant way. That is, above the noise level of instrumentation appropriate to the task of measuring the diagnostic parameters. For each marker pre-selected according to these criteria, a matrix is established in which columns represent samples, rows represent markers and each element is a normalized intensity measurement for the expression of that marker according to the relationship:

55

$$\left| \frac{(\mu_t - I)}{\mu_t} \right|$$

where  $I$  is the intensity measurement.

[0029] It is also possible to set additional boundary conditions to define the optimal portfolios. For example, portfolio size can be limited to a fixed range or number of markers. This can be done either by making data pre-selection criteria more stringent (e.g.,

$$.8 \leq \left| \frac{(\mu_t - MAX_n)}{(\sigma_t + \sigma_n)} \right|$$

10 instead of

$$0.5 \leq \left| \frac{(\mu_t - MAX_n)}{(\sigma_t + \sigma_n)} \right|$$

15

or by using programming features such as restricting portfolio size. One could, for example, set the boundary condition that the efficient frontier is to be selected from among only the most optimal 10 genes. One could also use all of the genes pre-selected for determining the efficient frontier and then limit the number of genes selected (e.g., no more than 10).

20 [0030] The process of selecting a portfolio can also include the application of heuristic rules. Preferably, such rules are formulated based on biology and an understanding of the technology used to produce clinical results. More preferably, they are applied to output from the optimization method. For example, the mean variance method of portfolio selection can be applied to microarray data for a number of genes differentially expressed in subjects with breast cancer. Output from the method would be an optimized set of genes that could include some genes that are expressed

25 in peripheral blood as well as in diseased breast tissue. If sample used in the testing method are obtained from peripheral blood and certain genes differentially expressed in instances of breast cancer could also be differentially expressed in peripheral blood, then a heuristic rule can be applied in which a portfolio is selected from the efficient frontier excluding those that are differentially expressed in peripheral blood. Of course, the rule can be applied prior to the formation of the efficient frontier by, for example, applying the rule during data pre-selection.

30 [0031] Other heuristic rules can be applied that are not necessarily related to the biology in question. For example, one can apply the rule that only a given percentage of the portfolio can be represented by a particular gene or genes. Commercially available software such as the Wagner Software readily accommodates these types of heuristics. This can be useful, for example, when factors other than accuracy and precision (e.g., anticipated licensing fees) have an impact on the desirability of including one or more genes.

35 [0032] One method of the invention involves comparing gene expression profiles for various genes (or portfolios) to conduct diagnoses as described above. The gene expression profiles of each of the genes comprising the portfolio are fixed in a medium such as a computer readable medium. This can take a number of forms. For example, a table can be established into which the range of signals (e.g., intensity measurements) indicative of disease is input. Actual patient data can then be compared to the values in the table to determine whether the patient samples are normal or

40 diseased. In a more sophisticated embodiment, patterns of the expression signals (e.g., fluorescent intensity) are recorded digitally or graphically. The gene expression patterns from the gene portfolios used in conjunction with patient samples are then compared to the expression patterns. Pattern comparison software can then be used to determine whether the patient samples have a pattern indicative of the disease in question. Of course, these comparisons can also be used to determine whether the patient results are normal. The expression profiles of the samples are then

45 compared to the portfolio of a normal or control cell. If the sample expression patterns are consistent with the expression pattern for a colorectal cancer then (in the absence of countervailing medical considerations) the patient is diagnosed as positive for colorectal cancer. If the sample expression patterns are consistent with the expression pattern from the normal/control cell then the patient is diagnosed negative for colorectal cancer.

50 [0033] Numerous well known methods of pattern recognition are available. The following references provide some examples:

Weighted Voting: Golub, TR., Slonim, DK., Tamaya, P., Huard, C., Gaasenbeek, M., Mesirov, JP., Coller, H., Loh, L., Downing, JR., Caligiuri, MA., Bloomfield, CD., Lander, ES. *Molecular classification of cancer: class discovery and class prediction by gene expression monitoring*. Science 286:531-537, 1999

55 Support Vector Machines: Su, AI., Welsh, JB., Sapinosa, LM., Kern, SG., Dimitrov, P., Lapp, H., Schultz, PG., Powell, SM., Moskaluk, CA., Frierson, HF. Jr., Hampton, GM. *Molecular classification of human carcinomas by use of gene expression signatures*. Cancer Research 61:7388-93, 2001 and

Ramaswamy, S., Tamayo, P., Rifkin, R., Mukherjee, S., Yeang, CH., Angelo, M., Ladd, C., Reich, M., Latu-

lippe, E., Mesirov, J.P., Poggio, T., Gerald, W., Loda, M., Lander, E.S., Gould, T.R. *Multiclass cancer diagnosis using tumor gene expression signatures* Proceedings of the National Academy of Sciences of the USA 98:15149-15154, 2001

5 K-nearest Neighbors: Ramaswamy, S., Tamayo, P., Rifkin, R., Mukherjee, S., Yeang, C.H., Angelo, M., Ladd, C., Reich, M., Latulippe, E., Mesirov, J.P., Poggio, T., Gerald, W., Loda, M., Lander, E.S., Gould, T.R. *Multiclass cancer diagnosis using tumor gene expression signatures* Proceedings of the National Academy of Sciences of the USA 98:15149-15154, 2001

10 Correlation Coefficients: van 't Veer L.J., Dai H., van de Vijver M.J., He Y.D., Hart A.A., Mao M., Peterse H.L., van der Kooy K., Marton M.J., Witteveen A.T., Schreiber G.J., Kerkhoven R.M., Roberts C., Linsley P.S., Bernards R., Friend S.H. Gene expression profiling predicts clinical outcome of breast cancer. *Nature*. 2002 Jan 31;415(6871):530-6.

[0034] The gene expression profiles of this invention can also be used in conjunction with other non-genetic diagnostic methods useful in cancer diagnosis, prognosis, or treatment monitoring. For example, in some circumstances it is beneficial to combine the diagnostic power of the gene expression based methods described above with data from conventional markers such as serum protein markers (e.g., carcinoembryonic antigen). A range of such markers exists including such analytes as CA19-9, CA 125, CK-BB, and Guanylyl Cyclase C. In one such method, blood is periodically taken from a treated patient and then subjected to an enzyme immunoassay for one of the serum markers described above. When the concentration of the marker suggests the return of tumors or failure of therapy, a sample source amenable to gene expression analysis is taken. Where a suspicious mass exists, a fine needle aspirate is taken and gene expression profiles of cells taken from the mass are then analyzed as described above. Alternatively, tissue samples may be taken from areas adjacent to the tissue from which a tumor was previously removed. This approach can be particularly useful when other testing produces ambiguous.

[0035] Combining the use of genetic markers with other diagnostics is most preferred when the reliability of the other diagnostic is suspect. For example, it is known that serum levels of CEA can be substantially affected by factors having nothing to do with a patient's cancer status. It can be beneficial to conduct a combination gene expression/CEA assay when a patient being monitored following treatment for colon cancer shows heightened levels of routine CEA assays.

[0036] Articles of this invention include representations of the gene expression profiles useful for treating, diagnosing, prognosticating, and otherwise assessing diseases. These profile representations are reduced to a medium that can be automatically read by a machine such as computer readable media (magnetic, optical, and the like). The articles can also include instructions for assessing the gene expression profiles in such media. For example, the articles may comprise a CD ROM having computer instructions for comparing gene expression profiles of the portfolios of genes described above. The articles may also have gene expression profiles digitally recorded therein so that they may be compared with gene expression data from patient samples. Alternatively, the profiles can be recorded in different representational format. A graphical recordation is one such format. Clustering algorithms such as those incorporated in "GENESPRING" and "DISCOVER" computer programs mentioned above can best assist in the visualization of such data.

[0037] Different types of articles of manufacture according to the invention are media or formatted assays used to reveal gene expression profiles. These can comprise, for example, microarrays in which sequence complements or probes are affixed to a matrix to which the sequences indicative of the genes of interest combine creating a readable determinant of their presence. Alternatively, articles according to the invention can be fashioned into reagent kits for conducting hybridization, amplification, and signal generation indicative of the level of expression of the genes of interest for detecting colorectal cancer.

[0038] Kits made according to the invention include formatted assays for determining the gene expression profiles. These can include all or some of the materials needed to conduct the assays such as reagents and instructions.

45 [0039] The invention is further illustrated by the following non-limiting examples.

**Examples:** Genes analyzed according to this invention are identified by reference to Gene ID Numbers in the GenBank database. These are typically related to full-length nucleic acid sequences that code for the production of a protein or peptide. One skilled in the art will recognize that identification of full-length sequences is not necessary from an analytical point of view. That is, portions of the sequences or ESTs can be selected according to well-known principles for which probes can be designed to assess gene expression for the corresponding gene.

#### Example 1- Sample Handling and LCM.

55 [0040] Twenty-seven fresh frozen tissue samples were collected from patients who had surgery for a colorectal tumor. Nineteen of the samples were colorectal malignancy specimens, and eight of the samples were of normal colon mucosa. The tissues were snap frozen in liquid nitrogen within 20-30 minutes of harvesting, and stored at -80C° thereafter. For laser capture, the samples were cut (6µm), and one section was mounted on a glass slide, and the second on film (P. A.L.M.), which had been fixed onto a glass slide (Micro Slides Colorfrost, VWR Scientific, Media, PA). The section

mounted on a glass slide was after fixed in cold acetone, and stained with Mayer's Haematoxylin (Sigma, St. Louis, MO). A pathologist analyzed the samples for diagnosis and grade. The clinical stage was estimated from the accompanying surgical pathology and clinical reports, using the Dukes classification. The section mounted on film was after fixed for five minutes in 100% ethanol, counter stained for 1 minute in eosin/100% ethanol (100µg of Eosin in 100ml of dehydrated ethanol), quickly soaked once in 100% ethanol to remove the free stain, and air dried for 10 minutes.

[0041] Two of the colorectal adenocarcinomas were of grade 1, 10 of grade 2, and 5 of grade 3. One of the malignant samples was a carcinoid tumor of the caecum, and one a metastatic melanoma lesion. Two of the adenocarcinoma samples represented the mucinous subtype, and one the signet cell subtype. The Dukes staging of the adenocarcinomas divided them as follows: Dukes A: 2, Dukes B: 5, Dukes C: 7, Dukes D: 3. Six of the adenocarcinomas had been irradiated preoperatively.

[0042] Before use in LCM, the membrane (LPC-MEMBRANE PEN FOIL 1.35µm No 8100, P.A.L.M. GmbH Mikrolaser Technologie, Bernried, Germany) and slides were pretreated to abolish RNases, and to enhance the attachment of the tissue sample onto the film. Briefly, the slides were washed in DEP H<sub>2</sub>O, and the film was washed in RNase AWAY (Molecular Bioproducts, Inc., San Diego, CA) and rinsed in DEP H<sub>2</sub>O. After attaching the film onto the glass slides, the slides were baked at +120°C for 8 hours, treated with TI-SAD (Diagnostic Products Corporation, Los Angeles, CA, 1:50 in DEP H<sub>2</sub>O, filtered through cotton wool), and incubated at +37°C for 30 minutes. Immediately before use, a 10µl aliquot of RNase inhibitor solution (Rnasin Inhibitor 2500U=33U/µl N211A, Promega GmbH, Mannheim, Germany, 0.5µl in 400µl of freezing solution, containing 0.15 mol NaCl, 10 mmol Tris pH 8.0, 0.25 mmol dithiothreitol) was spread onto the film, where the tissue sample was to be mounted.

[0043] The tissue sections mounted on film were used for LCM. Approximately 2000 epithelial cells/sample were captured using the PALM Robot-Microbeam technology (P.A.L.M. Mikrolaser Technologie, Carl Zeiss, Inc., Thornwood, NY), coupled into Zeiss Axiovert 135 microscope (Carl Zeiss Jena GmbH, Jena, Germany). The surrounding stroma in the normal mucosa, and the occasional intervening stromal components in cancer samples, were included. The captured cells were put in tubes in 100% ethanol and preserved at -80°C.

#### **Example 2- RNA Extraction and Amplification.**

[0044] Zymo-Spin Column (Zymo Research, Orange, CA 92867) was used to extract total RNA from the LCM captured samples. About 2 ng of total RNA was resuspended in 10 µl of water and 2 rounds of the T7 RNA polymerase based amplification were performed to yield about 50 µg of amplified RNA.

#### **Example 3- cDNA Microarray Hybridization and Quantitation.**

[0045] A set of cDNA microarrays consisting of approximately 20,000 human cDNA clones was used to test the samples. About 30 plant genes were also printed on the microarrays as a control for non-specific hybridization. Cy3-labeled cDNA probes were synthesized from 5 µg of aRNA of the LCM captured cells. The probes were purified with Qiagen's Nucleotide Removal Columns and then hybridized to the microarrays for 14-16 hours. The slides were washed and air-dried before scanning. cDNA microarrays were scanned for cy3 fluorescence and ImaGene software (Biodiscovery, Los Angeles, CA) was used for quantitation. For each cDNA clone, four measurements were obtained using duplicate spots and duplicate arrays and the intensities were averaged.

[0046] cDNAs were printed on amino silane-coated slides (Coming) with a Generation III Micro-array Spotter (Molecular Dynamics). The cDNAs were PCR amplified, purified (Qiagen PCR purification kit), and mixed 1:1 with 10 M NaSCN printing buffer. Prior to hybridization micro-arrays were incubated in isopropanol at room temperature for 10 min. The probes were incubated at 95°C for 2 min, at room temperature for 5 min, and then applied to three replicate slides. Cover slips were sealed onto the slides with DPX (Fluka) and incubated at 42°C overnight. Slides were then washed at 55°C for 5min in 1X SSC/0.2% SDS and 0.1X SSC/0.2% SDS, dipped in 0.1X SSC and dried before being scanned by a GenIII Array Scanner (Molecular Dynamics). The fluorescence intensity for each spot was analyzed with AUTOGENE software (Biodiscovery, Los Angeles).

[0047] Chip intensities were linearly normalized forcing the intensity reading at the 75<sup>th</sup> percentile equivalent to a value of 100 on each chip. Every gene on the chip was normalized to itself by dividing the intensity reading for that gene by the median of the gene's expression value readings over all the samples. Prior to clustering, genes that did not have an intensity reading of 100 or greater in at least one sample were filtered out in order to limit the background affect on the similarity metrics. A set of 6,225 genes was selected for clustering analysis. Hierarchical clustering was performed using correlation as a measure of similarity, which groups together samples with genes that are showing positive changes at the same time without any consideration for negative changes (Silicon Genetics, Sunnyville, CA). Each of the major nodes in the dendrogram was then considered a subgroup of samples. Differentially expressed genes were identified by comparing each tumor subgroup to the normal group. The selection was based on a signal to noise measurement threshold given by



$$1 \leq \left| \frac{(\mu_t - \mu_n)}{(\sigma_t + \sigma_n)} \right|,$$

5 where  $\mu_t$  is the mean of the tumor subset,  $\mu_n$  is the mean of the subset of normal samples, and  $\sigma_t + \sigma_n$  represent the combined standard deviations. The within-group coefficient of variation of the intensity readings of a gene had to be less than 0.33, for the gene to be included in the pair-wise comparisons. The median of the tumor group over the median of the normal group had to be greater than, or equal to 2 for up-regulation, and less than, or equal to 0.5 for down-regulation. If a gene met all the criteria, it was selected. The genes selected in all the comparisons were considered consistently dysregulated in colorectal cancer. The p-values for the statistical significance were calculated using a T-test assuming unequal variance. The gene set for clustering was also subjected to principal component analysis (PCA) using a software package (Partek, St Louis, MO). The data was then projected onto the reduced 3-dimensional space. The normal and tumor colorectal samples were represented by the projected expression levels.

10 **[0048]** A list of genes with large up-regulated differentials was created to distinguish between the tumor and normal samples. One-hundred and twenty-three genes were preselected by using

$$0.5 \leq \left| \frac{(\mu_t - MAX_n)}{(\sigma_t + \sigma_n)} \right|$$

20

as a signal to noise cutoff. A ratio equal to, or greater than 1.5 was the minimal criterion for up-regulation. Genes were also included if

$$0.9 \leq \left| \frac{(\mu_t - \mu_n)}{(\sigma_t + \sigma_n)} \right|.$$

A portfolio of four genes was established, each having at least a three fold expression differential between tumor and normal cells.

30 **[0049]** Differentially Expressed Genes in Colorectal Cancer. Thirty-nine genes were differentially expressed in all tumor samples as compared to normal colon mucosa. Thirty-seven of them were significantly down-regulated in all the tumors, except for an outlier. Two of them were up-regulated. The identities of the genes were verified by sequencing the cDNA clones placed on the microarray. Results are shown in Table 1.

35

Table 1

Modulated Genes					
ACCESSION	GENE DESCRIPTION	MEAN SIGNAL INTENSITY (NORMAL)	MEAN SIGNAL INTENSITY (TUMOR)	P-VALUE	
AF071569	CaM kinase II gene subtype delta 2.	93	39	4.64E-09	Seq. ID No. 1
AB014530	Homo sapiens mRNA for KIAA0630 protein	108	50	4.83E-07	Seq. ID No.2
AK000319	Human cDNA KIAA0630	236	69	7.84E-06	Seq. ID No.3
U81504	beta-3A-adaptin subunit of the AP-3 complex mRNA,	241	75	3.52E-05	Seq. ID No.4
AB011166	Human cDNA KIAA0594	116	55	3.53E-05	Seq. ID No.5
AB040914	Human cDNA KIAA1481	187	59	8.85E-05	Seq. ID No.6

Table 1 (continued)

Modulated Genes					
ACCESSION	GENE DESCRIPTION	MEAN SIGNAL INTENSITY (NORMAL)	MEAN SIGNAL INTENSITY (TUMOR)	P-VALUE	
AK025205	Human cDNA FLJ21552	322	97	0.00013	Seq. ID No.7
AJ278219	Fatty acid hydroxylase	143	53	0.00011	Seq. ID No.8
AB046854	Human cDNA KIAA1634	142	59	0.00020	Seq. ID No.9
R00585	Unknown	149	57	1.28E-09	Seq. ID No.10
S45844	Spi-B transcription factor	140	43	0.00043	Seq. ID No.11
X98311	Carcinoembryonic antigen family member 2 (CGM2)	6137	223	0.00044	Seq. ID No.12
BAA78050	NADPH oxidoreductase homolog	153	84	0.00048	Seq. ID No.40
N72128	Unknown	164	77	0.00068	Seq. ID No.13
AB040955	Human cDNA KIAA1552	334	120	0.00067	Seq. ID No.14
AF125101	HSPC040 protein	363	115	0.0011	Seq. ID No.15
AB023229	Human cDNA KIAA1012	263	88	0.00099	Seq. ID No.16
N95761	$\alpha$ -L-fucosidase gene	429	104	0.00047	Seq. ID No.17
AK025033	Human cDNA FLJ21380	180	85	0.0010	Seq. ID No.18
L10844	Human cellular growth regulating protein	206	101	0.0013	Seq. ID No.19
H96534	H.sapiens mRNA for gp25L2 protein.	147	58	0.0015	Seq. ID No.20
AK001521	Human cDNA FLJ10659	157	60	0.0019	Seq. ID No.21
AF151039	HSPC205 protein	117	60	0.0017	Seq. ID No.22
AF052059	SEL 1L protein	168	53	0.0016	Seq. ID No.23
N24597	Unknown	166	62	0.0016	Seq. ID No.24
AK001950	Inner centromere protein	148	64	0.0029	Seq. ID No.25
BAA02649	Macrophage scavenger receptor type I	118	44	0.0031	Seq. ID No.41
N75004	Unknown	98	48	0.0031	Seq. ID No.26

Table 1 (continued)

Modulated Genes					
ACCESSION	GENE DESCRIPTION	MEAN SIGNAL INTENSITY (NORMAL)	MEAN SIGNAL INTENSITY (TUMOR)	P-VALUE	
W16916	Human cDNA KIAA0260	162	61	0.0037	Seq. ID No.27
X52001	H.sapiens endothelin 3 mRNA.	89	33	0.0042	Seq. ID No.28
T50788	Unknown	364	102	0.0059	Seq. ID No.38
AJ005866	Putative Sqv-7 like protein	381	163	0.0049	Seq. ID No.29
AF113535	MAID protein	218	100	0.0053	Seq. ID No.39
AB037789	Human cDNA KIAA1368	164	62	0.0068	Seq. ID No.30
M33987	Carbonic anhydrase	652	46	0.0074	Seq. ID No.31
M77830	Desmoplakin 1 (DPI)	184	81	0.0092	Seq. ID No.32
H81220	EST domain transcription factor ELF1	113	55	0.017	Seq. ID No.33
AF000592	Human chromosome 21q11-q21 genomic clone	33	69	1.16E-05	Seq. ID No.35
AK021701	Human cDNA FLJ11639	31	63	0.00070	Seq. ID No.36

**Example 4: Optimized Portfolio for Colorectal Tumors.**

[0050] The mean-variance optimization algorithm was used to generate a multiple genebased signature, where the genes that are included can be used in combination to distinguish between the normal and tumor samples. Intensity measurements were processed using the samples and microarrays described in Examples 1-3. The data to be analyzed was first preselected based on a pre-specified 5-fold differential between tumor and normal cells. The expression data from genes preselected according to this criteria were then used as follows. The mean and standard deviation of the intensity measurements for each gene were calculated using the non-metastatic samples as the baseline. A discriminating value of  $X \cdot (\text{Standard Deviation} + \text{Mean})$  was then calculated for each baseline gene (X was assigned a value of 3). This value was used to ensure the resulting portfolio would be stringent. A ratio of the discriminating value to the baseline value was then calculated for each metastatic sample. This ratio was then converted to a common logarithm. This data was then imported into Wagner Software which produced an efficient frontier from which a portfolio of 4 genes was selected. The set included an unknown sequence, procollagen type I, large subunit of ribosomal protein L21 and fibronectin. These genes are identified as Seq. ID No. 42, Seq. ID No. 43, Seq. ID No. 44 and Seq. ID No. 45. Alternatively, a combination of genes used to make up the portfolio can be used to produce diagnostic information that is useful for making clinical decisions regarding colorectal cancer. This is particularly beneficial in the case when a combination of genes selected from the portfolio are combined with additional markers (genetic or not).

Optimized Gene Portfolio:

**[0051]**

- 5 >gil1264443|gb|IN92134.1|IN92134 za23f09.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE: 293417 5' similar to gb|IM87908|HUMALNE32 Human carcinoma cell-derived Alu RNA transcript, (rRNA); gb: X57025\_mal INSULIN-LIKE GROWTH FACTOR IA PRECURSOR (HUMAN)
- >gil2221047|gb|AA490172.1|AA490172 ab06b08.s1 Stratagene fetal retina 937202 Homo sapiens cDNA clone IMAGE:839991 3' similar to gb:J03464 PROCOLLAGEN ALPHA 2(I) CHAIN PRECURSOR (HUMAN)
- 10 >gil2188918|gb|AA464034.1|AA464034 zx86b09.r1 Soares ovary tumor NbHOT Homo sapiens cDNA clone IMAGE:810617 5' similar to SW:RL21\_HUMAN P46778 60S RIBOSOMAL PROTEIN L21.
- >gil834491|gb|R62612.1|R62612 yi12d01.s1 Soares placenta Nb2HP Homo sapiens cDNA clone IMAGE:139009 3' similar to gb:X02761\_cds1 FIBRONECTIN PRECURSOR (HUMAN);

- 15 **[0052]** Using a different set of criteria but the same method, a further four gene portfolio was selected by the software. These are Seq. ID no. 46, Seq. ID No. 47, Seq. ID No. 48 and Seq. ID No. 49. Two genes overlap with the first four-gene portfolio. The two optimized portfolios can also be combined to form a six-gene portfolio.

Optimized Gene Portfolio:

20

**[0053]**

- >gil2114953|gb|AA431245.1|AA431245 zw78d06.r1 Soares\_testis\_NHT Homo sapiens cDNA clone IMAGE: 782315 5' similar to WP:F36H1.2 CE05814 ANKYRIN LIKE
- 25 >gil2156172|gb|AA443497.1|AA443497 zw34d03.r1 Soares ovary tumor NbHOT Homo sapiens cDNA clone IMAGE:771173
- >gil2221047|gb|AA490172 ab06b08.s1 Stratagene fetal retina 937202 Homo sapiens cDNA clone IMAGE: 839991 3' similar to gb:J03464 PROCOLLAGEN ALPHA 2(I) CHAIN PRECURSOR (HUMAN)
- >gil1264443|gb|IN92134.1|IN92134 za23f09.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE: 293417 5' similar to gb|IM87908|HUMALNE32 Human carcinoma cell-derived Alu RNA transcript, (rRNA); gb: X57025\_mal INSULIN-LIKE GROWTH FACTOR IA PRECURSOR (HUMAN);
- 30

35

40

45

50

55

## SEQUENCE LISTING

<110> WANG, YIXIN  
 5 <120> COLORECTAL CANCER DIAGNOSTICS  
 <130> CDS 267 US NP  
 <140> TBD  
 10 <141> 2003-03-21  
 <150> 60/368,798  
 <151> 2002-03-29  
 <160> 49  
 15 <170> PatentIn version 3.1  
 <210> 1  
 <211> 1500  
 <212> DNA  
 20 <213> human  
 <400> 1  
 atggcttcga ccaccacctg caccaggctc acggacgagt atcagctttt cgaggagctt  
 60  
 ggaaaggggg cattctcagt ggtgagaaga tgtatgaaaa ttctactlgg acaaggatat  
 25 120  
 gctgccaaaa ttatcaacac caaaaagctt tctgctaggg atcatcagaa actagaaaga  
 180  
 gaagctagaa tctgcgtct tttgaagcac cctaattattg tgcgacttca tgatagcata  
 30 240  
 tcagaagagg gctttcacta cttggtgttt gatttagtta ctggaggatga actgtttgaa  
 300  
 gacatagtgg caagagaata ctacagtga gctgatgcca gtcattglat acagcagatt  
 35 360  
 ctagaaagtg ttaatcattg tcacctaaat ggcatagttc acagggacct gaagcctgag  
 420  
 aatttgcttt tagctagcaa atccaaggga gcagctgtga aattggcaga ctttggttta  
 40 480  
 gccatagaag ttcaagggga ccagcaggcg tggtttggtt ttgctggcac acctggatat  
 540  
 ctttctccag aagttttacg taaagatcct tatggaaaac cagtggatat gtgggcatgt  
 45 600  
 ggtgtcattc tctatattct acttggtggg tatccacctt tctgggatga agaccaaac  
 660  
 agactctatc agcagatcaa ggctggagct tatgattttc catcaccaga atgggacacg  
 50 720  
 gtgactcclg aagccaaaga cctcatcaat aaaatgctta ctatcaacct tgccaaacgc  
 780  
 atcacagcct cagaggcact gaagcaccca tggatctgtc aacgttctac tgttgcttcc  
 55 840

atgatgcaca gacaggagac tgtagactgc ttgaagaaat ttaatgctag aagaaaacta  
900

5 aagggtgccca tcttgacaac tatgtggct acaaggaatt tctcagcagc caagagtgtg  
960

ttgaagaaac cagatggagt aaaggagtca actgagagtt caaatacaac aaltgaggat  
1020

10 gaagatgtga aagcacgaaa gcaagagatt atcaaagtca ctgaacaact gatcgaagct  
1080

atcaacaatg gggactttga agcctacaca aaaatctgtg acccaggcct tactgctttt  
1140

15 gaacctgaag ctttgggtaa tttagtggaa gggatggatt ttcacagatt ctactttgaa  
1200

aatgctttgt ccaaaagcaa taaaccaatc cacactatta ttctaaacc tcatgtacat  
1260

20 ctggtagggg atgatgccgc ctgcatagca tatattaggc tcacacagta catggatggc  
1320

agtggaatgc caaagacaat gcagtcagaa gagactcgtg tgtggcaccg ccgggatgga  
1380

25 aagtggcaga atgttcattt tcctgcctcg gggtcaccaa caglacccal caagccaccc  
1440

tgtattccaa atgggaaaga aaacttctca ggaggcacct ctttgtggca aaacatctga  
1500

30 <210> 2  
<211> 5761  
<212> DNA  
<213> human

35 <400> 2  
cacaccgag tatgcggtag cctttactct gagctgcgca gccggccggc cggcgtggt  
60

tgaacagact gccgtgtac tggcgtggcc tggagggact cagcaaatc tctgccttc  
120

40 aacttggcaa cagttgcctg gggtagctct acacaactct gtccagccca cagcaatgat  
180

tccagaggcc atggggagtg gacagcagct agctgactgg aggaatgcc actctcatgg  
240

45 caaccagtac agcactatca tgcagcagcc atccttgctg actaaccatg tgacattggc  
300

cactgctcag cctctgaatg ttggtgttgc ccatgttgc agacaacaac aatccagttc  
360

50 cctcccttcg aagaagaata agcagtcagc tccagtctct tccaagtcct ctctagatgt  
420

tctgccttcc caagtctatt ctctggttgg gagcagtcct ctcgcacca catcttctta  
480

55 taattccttg gtcctgtcc aagatcagca tcagcccatc atcaltccag atactccag  
540

ccctcctgtg agtgtcatca ctatccgaag tgacactgat gaggaagagg acaacaaata  
600

5 caagcccagt agctctggac tgaagccaag gtctaattgc atcagttatg tcaactgtcaa  
660

tgattctcca gactctgact cttctttgag cagcccttat tccactgata ccttgagtgc  
720

10 tctccgaggc aatagtggat ccgttttgga ggggcctggc agagtgtgg cagatggcac  
780

tggcacccgc actatcattg tgccctcact gaaaactcag cttggtgact gcaactgtagc  
840

15 aaccagggc tcaggtctcc tgagcaataa gactaagcca gtcgcttcag tgagtgggca  
900

gtcatctgga tgctgtatca cccccacagg gtatcgagct caacgcgggg ggaccagtgc  
960

20 agcacaacca ctcaatctta gccagaacca gcagtcacg gcggctcaa cctcacagga  
1020

gagaagcagc aaccagccc cccgcaggca gcaggcggtt gtggccctc tctcccaagc  
1080

25 cccctacacc ttccagcatg gcagcccgct aactcgaca gggcaccac accttgcccc  
1140

ggccctgct cactgcca gccaggtca tctgtatacg tatgctgcc cgacttctgc  
1200

30 tgctgcactg ggctcaacca gctccattgc tcatctttc tcccacagg gttcctcaag  
1260

gcattgtgca gctatacca ctaccctag cactttggtg caccaggtcc ctgtcagtg  
1320

35 tgggcccagc ctccctcact ctgccagcgt ggccctgct cagtaccaac accagtttgc  
1380

caccaatcc tacattgggt cttcccgagg ctcaacaatt tacactggat acccgctgag  
1440

40 tcctaccaag atcagccagt attcctact atagtgggtg agcatgaggg aggaggaatc  
1500

atggctacct tctcctggcc ctgcgttctt aatattgggc tatggagaga tctccttta  
1560

45 cctcttgaa atttcttagc cagcaactg ttctgcagg gccactgaa gcaqaaggtt  
1620

tttctctggg ggaacctgct tcagtgttga ctgcattgtt gtaglcttcc caaagtttgc  
1680

50 cctattttta aattcattat tttgtgaca glaattttg tacttggag agttcagatg  
1740

cccatcttct gcagttacca aggaagagag attgttctga agttaccctc tgaaaaatat  
1800

55 tttgtctctc tgacttgatt tctataaatg cttttaaaaa caagtgaagc cctctttat  
1860

EP 1 355 149 A2

ttcattttgt gttattgtga ttgctggcca ggaaaaatgc tgatagaagg agllgaaatc  
1920

5 tgatgacaaa aaaagaaaaa ttactttttg ttgtttata aactcagact tgcctatttt  
1980

attttaaaag cggcttacac aatctccctt ttgtttattg gacatttaaa cttacagagt  
2040

10 ttcagttttg ttttaaatgc atattatact taatgggcaa ttgttatttt tgcaaaaactg  
2100

gttacgtatt actctgtgtt actattgaga ttctctcaat tgctcctgtg ttgtttataa  
2160

15 agtagtggtt aaaaggcagc tcaccatttg ctggtaactt aatgtgagag aatccatatt  
2220

tgcttgaaaa caccaagtat tctttttaaa tgaagcacca tgaattcttt tttaaattat  
2280

20 tttttaaaag tctttctctc tctgattcag cttaaatttt tttatcgaaa aagccattaa  
2340

ggtgggtatt attacatggt ggtgggtggt ttattatatg caaaatctct gtctattatg  
2400

25 agatactggc attgatgagc ttgacctaaa gatlagtatg aattttcagt aatacacctc  
2460

tgttttgctc atctctccct tctgttttat gtgatttgtt tggggagaaa gctaaaaaaa  
2520

30 cctgaaacca gataagaaca tttcttgtgt atagctttta tacttcaaag tagcttcctt  
2580

tgatgccag cagcaaattg aatgctctct tattaagact tatataataa gtgcatgtag  
2640

35 gaattgcaaa aaatatttta aaaatttall actgaattta aaaatatttt agaagttttg  
2700

taatgggtgt gttttaatat ttacataat taaatatgta catattgatt agaaaaatat  
2760

40 aacaagcaat ttttcctgct aacccaaaat gttatttgta atcaaattgt tagtgattac  
2820

acttgaattg tgtacttagt gtgtatgtga tcctccagtg ttatcccgga gatggattga  
2880

45 tgtctccatt gtatttaaac caaaatgaac tgatacttgt tggaatgtat gtgaactaat  
2940

tgcaattata ttagagcata ttactgtagt gctgaatgag caggggcatt gcctgcaagg  
3000

50 agaggagacc cttggaattg ttttgacacag gtgtgtctgg tgaggagttt ttcagtgtgt  
3060

gtctcttctt tccctttctt cctccttccc ttattgtagt gccttatatg ataattgtagt  
3120

55 ggttaataga gtttacagtg agcttgccctt aggatggacc agcaagcccc cgtggacctt  
3180



EP 1 355 149 A2

aagttgttca ccgggattta tcagaacagg attagtagct gtattgtgta atgcattgtt  
3240

5 ctcagtttcc ctgccaacat tgaaaaataa aaacagcagc ttttctcctt taccaccacc  
3300

tttaccctt tccatttttg attctcggcl gagttctcac agaagcattt tccccatgtg  
3360

10 gctctctcac tgtgcgttgc taccttgctt ctgtgagaat tcaggaagca ggtgagagg  
3420

gtcaagccaa tattaaatat gcattctttt aaagtatgtg caatcacttt tagaatgaat  
3480

15 ttttttttcc ttttcccatg tggcagtcct tcctgcacat agttgacatt cctagtaaaa  
3540

tatttgcttg ttgaaaaaaa catgttaaca gatgtgttta taccaaagag cctgttglat  
3600

20 tgcttaccat gtccccatac tatgaggaga agttttgtgg tgccgctggg gacaaggaac  
3660

tcacagaaaag gtttcttagc tggatgaagaa tatagagaag gaaccaaagc ctgttgagtc  
3720

25 attgaggctt ttgaggtttc ttttttaaca gcttgatatag tcttggggcc cttcaagctg  
3780

tgaaattgtc cttgtactct cagctcctgc atggatctgg gtcaagtaga aggtactggg  
3840

30 gatggggaca ttctgcccc taaaggattt ggggaagaa gattaatcct aaaatacagg  
3900

tgtgttccat ccgaattgaa aatgatatat ttgagatata attttaggac tggttctgtg  
3960

35 tagatagaga tgggtgcaag gaggtgcagg atggagatgg gagatttcat ggagcctgg  
4020

cagccagctc tgtaccaggt tgaacaccga ggagctgtca aagtatttgg agtttcttca  
4080

40 ttgtaaggag taagggttc caagatgggg caggtagtcc gtacagccta ccaggaacat  
4140

gttgtgtttt ctttattttt taaaatcatt atattgagtt gtgttttcag cactatattg  
4200

45 gtcaagatag ccaagcagtt tgtataattt ctgtcactag tgtcatacag ttttctggtc  
4260

aacatgtgtg atctttgtgt ctcttttttg ccaagcacat tctgatlttc ttgttgaac  
4320

50 acaggtctag tttctaaagg acaaattttt tgttccttgt cttttttctg taagggacaa  
4380

gatttgttgt ttttgtaaga aatgagatgc aggaaagaaa accaaatccc attcctgcac  
4440

55 cccagtccaa taagcagata ccacttaaga taggaqtcta aactccacag aaaaggataa  
4500

taccaagagc ttgtattggt accttagtca cttgcctagc agtgtgtggc tttaaaaact  
 4560  
 5 agagatTTTT cagtcttagt ctgcaaacgt gcatttccga tttccagca taaaaatcca  
 4620  
 cctgtgtctg ctgaatgtgt atgtatgtgc tcaactgtggc tttagattct gtccctgggg  
 4680  
 10 ttagccctgt tggccctgac aggaaggag gaagcctggt gaatttagtg agcagctggc  
 4740  
 ctgggtcaca gtgacctgac ctcaaaccag ctttaaggctt taagtccctc ctcagaactt  
 4800  
 15 ggcatttcca acttcttctt ttcgggtga gagaagaagc ggagaagggt tcagtgtagc  
 4860  
 cactctgggc tcataggagc acttggtcac tccagagttt ttaatagctc ccaggagggtg  
 4920  
 20 atattatTTT cagtgtctag ctgaaatacc aaccacagga ataagaactc catttcaaac  
 4980  
 agttctggcc attctgagcc tgcttttgtg attgtctatc cattgtcttc cactagaggg  
 5040  
 25 gctaagcttg actgccctta gccaggcaag cacagtaatg tgtgttttgt tcagcattat  
 5100  
 tatgcaaaaa ttcactagtt gagalggttt gtttttaggt aggaaatgaa attgcctctc  
 5160  
 30 agtgacagga gtggcccgag cctgcttctt attttgattt tttttttttt taactgatag  
 5220  
 atggtgcagc atgtctacat ggttgtttgt tgctaaactt tatataatgt gtggtttcaa  
 5280  
 35 ttcagcttga aaaataatct cactacatgt agcagtacat tatatgtaca ttatatgtaa  
 5340  
 tggttagtatt tctgcttga atccttgata ttgcaatgga attcctactt tattaaatgt  
 5400  
 40 atttgatatg ctagtatttg tgtgcgattt aaactttttt tgctttctcc ctttttttgg  
 5460  
 ttgtgcgctt tctttttacaa caagcctcta gaaacagata gtttctgaga attactgagc  
 5520  
 45 tatgtttgta atgcagatgt acttagggag tatgtaaaat aatcatttta acaaaagaaa  
 5580  
 tagatattta aaatttaata ctaactatgg gaaaagggtc cattgtgtaa aacatagttt  
 5640  
 50 atctttggat tcaatgtttg tctttggttt tacaaagtag cttgtatttt cagtattttc  
 5700  
 tacataatat ggtaaaatgt agagcaattg caatgcatca ataaaatggg taaattttct  
 5760  
 55 g  
 761

EP 1 355 149 A2

5 <210> 3  
 <211> 2129  
 <212> DNA  
 <213> human  
  
 <400> 3  
 ctgtattgag acaaaggaag ggatctgtca gaaagcaaca ctgttatct tgggcttggc  
 60  
 10 agcaaggaag aggacaggta gtggagatcc tgcaatctga aaagcagact gaaagggtgac  
 120  
 aaagaagctg aagatgggtg gtggagagag gtataacatt ccagcccctc aatctagaaa  
 180  
 15 tgtttagtaag aaccaacaac agcttaacag acagaagacc aaggaacaga attcccagat  
 240  
 gaagattgtt cataagaaaa aagaagagg acatgggtat aactcatcag cagctgcctg  
 300  
 20 gcaggccatg caaatgggg ggaagaacaa aaattttcca aataatcaaa gttggaatcc  
 360  
 tagcttatca ggtcccaggt tactttttta atctcaagct aatcagaact atgctggtgc  
 420  
 25 caaatttagt gagccgccat caccaagtgt tcttcccaaa ccaccaagcc actgggtccc  
 480  
 tgtttccttt aatccttcag ataaggaaat aatgacattt caacttaaaa ccttacttaa  
 540  
 30 agtacaggta taaaataaga caaatgttta aatlttagta tgttcacggg tagttgtcaa  
 600  
 ttggtctgaa acaaatttgc tagggaatct atttgtgtag aactaattaa tgtaaaaaaa  
 660  
 35 atagaccatc tcgtgttggtg tgcactgtga tataatggta gtatcagtgc aacttaaact  
 720  
 aatgattgta cttgatatta agtgttctca actgagtaac ttttaagtgg aaaccaagtt  
 780  
 40 tagatttggg gagtggtaaa ggaatcagct ttttctattg ttaggggaag acagtaattt  
 840  
 atcattcatg gaccagtaga ttgttgaaag ttggtgaatc ggattataag cttctagcta  
 900  
 45 acacaaggat tcagaattag gtaaacatct gaaggtttag tatattagaa acacccaaac  
 960  
 cagtaatatg ctaacctgat gcactgctga aagaaaatgt gaatttttcg taataattgc  
 1020  
 50 atttttagta attgtacagt ggggtgaaag ggcatttgga gctcattaga atgagacata  
 1080  
 gtacaccca atggccctgt ttattaaatg tagtggatta agtgtctgtc aacaaataca  
 1140  
 55 ccaaaccat ttttataga aacagtattt aatggtcact caatagcttt caaaatacat

1200

5      ttttgtatta cagcactgca caagctattc taatagtgcct ctggcctcat cattcctgca  
1260

         aagccttgctt tggggagttg galaalgta aaattttaag tacctagggg agaaagagcc  
1320

10      atgtaaatat ctgtaataaa cttgtagcat atgtaaagt. ttcttggcct ttatcttaca  
1380

         aaaatggagt attttagtat gaatttgctg aatgtaagac cgtggactgt tttttataat  
1440

15      atggcctaata tttaaaggta caaaataact tgttttttaa gtttgcctt gtgctaaagt  
1500

         gccagtgtat gtatgttata cttgatttgg ttgtaaaacta tttttcaaag taaaccctag  
1560

20      tgtaataagt ttataacta aaaagggtta agctgctaaa actattttta agagatgtga  
1620

         aatgcagtat gggactatct ttttttctc ctctaagccc aaagattaac tagagtcctt  
1680

25      ccaaccctat agattgttgg ctttcacaat cttataacct aggatacagg lagtttcgag  
1740

         tatggtgccca gtgatgtttt gtttttgttt ggtcaagggg taggtgcaac ccaatggacc  
1800

30      acttatgcaa aagatgtaaa ctcttgcata atacattgat aacatgtttt gccaaactta  
1860

         aatgcttaaa cataagcgaa accagtagca agtatgtggg tcagcttaaa aattttgatt  
1920

35      gttaatgccc ttttttctaa ttggcacct cttgatgcct aagcaggtaa gcagatgcct  
1980

         aagctgtatt tctccaaata aatcaagatg aagtactgcc caagttaaatt attgatagcc  
2040

40      taaagacaag tttatgtagt acttaatgta catgatatga agcataaaat taaataaaat  
2100

         ttttcccat tgaaaaaaaa aaaaaaaaaa  
2129

45      <210> 4  
         <211> 3950  
         <212> DNA  
         <213> human

50      <400> 4  
         cgagaactag ttttgttccg tgccctctgg actggaacct tttggagaga acccccggca  
60

         ggaccaaccc cgcaccgcc agcaccgcc caatgtccag caatagtttt ccttacaatg  
120

55      agcagtcagg aggaggggag gcgacggagc tgggtcagga ggcgacctca accatttccc  
180

EP 1 355 149 A2

cctcgggggc cttcggcctc tttagcagcg atttgaagaa gaatgaagat ctlaagcaaa  
240

5 tgltagagag caacaaagat tctgctaaac tggatgctat gaagcggatt gttgggatga  
300

ttgcaaaagg gaaaaatgca tctgaactgt ttctgctgt tgtgaagaat gtggccagta  
360

10 aaaaatttga gatcaagaag ttggtatatg ttacctggt tcgatatgct gaagaacagc  
420

aggatcttgc actcctgtcc ataagcactt ttcagcgagc tctgaaggac ccaaaccaac  
480

15 taattcgtgc aagcgctttg agagtctgt caagtattag agtgccaatt attgtaccta  
540

tcatgatgct tgctattaag gaagcttctg ctgacttacc accatatgtt aggaagaatg  
600

20 cagcccatgc aatacaaaaa ttatacagcc ttgatccaga gcagaaggaa atgttaattg  
660

aagtaattga aaaacttctg aaagataaaa gcacattggt agctgqcagt gttgtgatgy  
720

25 cttttgaaga agtatgcccc gacagaatag atctgattca taaaaattac cgcaagctat  
780

gtaacttact agtggatgtt gaagagtggg ggcaggttgt cataatccac atgctaactc  
840

30 gatatgctcg gacacagttt gtcagccctt ggaaagaggg tgatgaatta gaagacaatg  
900

gaaagaattt ctacgaatct gatgatgac agaaggaaaa gactgacaaa aagaagaagc  
960

35 cgtatactat ggatccagat catagactct taattagaaa tacaagcct ttgcttcaga  
1020

gcaggaatgc tgcggtggtt atggcagttg ctcagctgta ttggcacata tcacaaaat  
1080

40 ctgaagctgg cataatttct aaatcactag tgcgtttact tcgtagcaat agggaggtgc  
1140

agtatatgt cctacaaaat atagcaacta tgtcaattca aagaaagggg atgtttgaac  
1200

45 cttatctgaa gagtttctat gttaggtcaa ctgatccaac tatgatcaag aactgaagc  
1260

ttgaaatttt gacaaacttg gcaaatgaag ccaacatalc aactcttctt cgagaatttc  
1320

50 agacctatgt gaaaagccag gataaacaat ttgcagcagc cactattcag actataggca  
1380

gatgtgcaac caacatcttg gaagtcactg acacgtgcct caatggcttg gtctgtctgc  
1440

55 tgtccaacag ggatgaaata gttgttgctg aaagtgtggt tgttataaag aaattactgc  
1500

EP 1 355 149 A2

aaatgcaacc tgcacaacat ggtgaaatta tlaaacatat ggccaaactc ctggacagta  
 1560  
 5 tcaactgttcc tgttgctaga gcaagtattc tttggctaata tggagaaaac tgtgaacgag  
 1620  
 ttcctaaaaat tgccctgat gttttgagga agatggctaa aagcttcact agtgaagatg  
 1680  
 10 atctggtaaa actgcagata ttaaatctgg gagcaaaatt gtatttaacc aactccaaac  
 1740  
 agacaaaatt gcttaccag tacatattaa atctcgga gatatgata aactacgaca  
 1800  
 15 tcagagaccg tacaagattt attagggcgc ttattgttcc gaatgaaaag agtggagctt  
 1860  
 taagtaata tgccaaaaa atattcctag caaaaagcc tgcaccactg cttgagtctc  
 1920  
 20 cttttaaga tagagatcat ttccagcttg gcaccttacc tcatactctc aacattaaag  
 1980  
 ctactggga cctggaatta tctaattggc cagaggtggc gcccgacca tcagttcgaa  
 2040  
 25 atgtagaagt aatagagttg gcaaaagaat ggaccccgagc aggaaaagca aagcaagaga  
 2100  
 attctgctaa gaagttttat tctgaatctg aggaagagga ggactcttct gatagtagca  
 2160  
 30 gtgacagtga gagtgaatct ggaagtgaat gtggagaaca aggcgaaagt ggggaggaag  
 2220  
 gagacagcaa tgaggacagc agtgaggact cctccagtga gcaggacagt gagaagtggac  
 2280  
 35 gggagtcagg cctagaaaac aaaagaacag ccaagaggaa ctcaaaagcc aaaggaaaaa  
 2340  
 gtgattctga agatggggag aaggaaaatg aaaaatctaa aacttcagat tcttcaaatg  
 2400  
 40 acgaatctag ttcaatagaa gacagttctt ccgattctga atcagagtca gaacctgaaa  
 2460  
 gtgaatctga atccagaaga gtcactaagg agaaagaaaa gaaaacaaag caagatagaa  
 2520  
 45 ctctctttac caaagatgtt tcaattctag atctggatga ttttaaccca gtatccactc  
 2580  
 cagttgcact tccacacca gctctttctc caagtttgat ggcagatctt gaaggtttac  
 2640  
 50 acttgtaac ttctcttca gtcactcagtg tcagtactcc tgcatctgta ccaacgaaaa  
 2700  
 ctacgtgct gcttcacga atgagtggaa aaggactagc tgccattat tctttccaa  
 2760  
 55 gacagccttg catTTTTTgt gataagatgg tctctatata aataacactg aataacacta  
 2820

# EP 1 355 149 A2

ctgatcgaaa gatagaaaat atccacatag gggaaaaaaa acttcctata ggcattgaaa  
2880

5 tgcattgttt taatccaata gactctcttg agcctgaggg atccattaca gtttcaatgg  
2940

gtattgactt ttgtgattct actcagactg ccagtttcca gttgtgtacc aaggatgatt  
3000

10 gcttcaatgt taatattcag ccacctgttg gagaactgct ttacctgtg gccatgtcag  
3060

agaaagattt taagaaagag caaggagtgc taacaggaat gaatgaaact tctgctgtaa  
3120

15 tcattgctgc accacagaat ttcactccct ctgtgatctt tcagaagggt gtaaattgtag  
3180

ccaatgtagg tgcagtcctt tctggccagg ataataaca caggtttgca gctaaaactg  
3240

20 tgcacagtgg gtcatgtatg ctatgcacag tggaaactgaa ggaaggctct acagcccagc  
3300

ttatcataaa cactgagaaa actgtgattg gctctgttcl gctgcgggaa ctgaagcctg  
3360

25 tctgtctca ggggtaacct gcttacatct ggactttaga atctggcaca caacaaaagt  
3420

gcctggcacc cactactgct gcctttcatt tataataata gcccttccat ctggcagtgg  
3480

30 gggtagaata cactcttgac attcttgtct cctgctttag aatgctagtg tgtatctatc  
3540

atgtatgcaa tactttcccc ctttttgctt tgctaaccga agagcatata ttttactgtc  
3600

35 agttgtctca actcttgaat ccattgtggc tttctctctg cctgctgctt cttttggcct  
3660

cctcgttttc cttctctttt tcgacaatgg tagacatgaa tgagatattt aaagttcatt  
3720

40 ggaaatcttc ttccctacag cagtaagcaa aaattagcaa agagatagtc taaatggcct  
3780

ctcagcttgg tatgtgaaaa tgagatcaca tactttttaa atccaaatar aaaagcatag  
3840

45 tctctgcaag attttgttct ttgaatttct tgatattgta attgattatt gataactgtc  
3900

atcatgaaat tatctctcaa taataagata aataaactag catatgaatc  
3950

50 <210> 5  
<211> 5191  
<212> DNA  
<213> human

55 <400> 5  
gagaaagaaa aacagctcga gacctcatgc aaagagaaaa ctgagtatct acagaaaatg  
60

gttcagagga atgaaagata taaacaagat gtggagaggt tctatgaacg gaagcgacat  
120

5 ttagatttaa ttgagatgct tgaagcaaaa aggcacatggg tggaatatga aaatgttcgt  
180

caggaatatg aagaagtaaa actagttcgt gaccgagtga aggaagaggt cagaaaactt  
240

10 aaagaagggc agattcctat aacatgtcga attgaagaaa tggaaaacga gcgtcacaaat  
300

ttggaggctc gaatcaaaga aaaggcaaca gatattaagg aggcacatcga aaaatgcaaa  
360

15 cagaagcaag atgttataga aaggaaagat aaacatattg aggaacttca gcaggcttta  
420

atagtaaagc aaaatgaaga gcttgaccga cagaggagaa taggtaatac ccgcaaaatg  
480

20 atagaggatt tgcaaaatga actaaagacc acggaaaact gcgagaatct tcagcccag  
540

attgatgcca ttacaaatga tctgagacgg attcaggatg aaaaggcatt atgtgaaggc  
600

25 gaaataattg ataagcgaag agagagggaa actctagaga aggagaaaaa gagtgtggac  
660

gatcatattg tacgttttga caatcttatg aatcagaagg aagataagct aagacagaga  
720

30 ttccgtgaca cgtatgatgc tgttttatgg ctaagaaata acagagacaa atttaaacia  
780

agagtctgtg agcccataat gctcacgacg aatatgaaag ataataaaaa tgccaaatat  
840

35 attgaaaatc atattccatc aaatgactta agagcctttg tatttgaaag tcaagaagat  
900

atggagggtt tcttcaaaga gggtcgtgac aataaaaaat taagagtaaa tgcgttatt  
960

40 gctccaaga gtccatatgc agacaaagca ccttcaagat ctttgaatga acttaacaa  
1020

tacggatttt tctcttattt gagagaatta ttgatgcac ctgatcctgt aatgagttac  
1080

45 ctttgctgtc agtatcatat tcatgaagtt cctgtaggaa ctgaaaagac cagagaaaga  
1140

attgaacggg taatacaaga aaccgatta aaacagattt atacagcaga agaaaagtat  
1200

50 gtggtgaaaa cttcttttta ttcaaacaaa gttatttcta gtaacacatc tctaaaagta  
1260

gcgcagtttc tcaatgtcac tgtggaccta gagcagagaa gacacttaga agaacagcta  
1320

55 aaggaaattc atagaaaatt gcaagcagtg gattcagggt tgattgcctt acgtgaaaca  
1380



EP 1 355 149 A2

agcaaacatc tggagcacia agacaatgaa cttagacaaa agaagaagga gcttcttgag  
1440

5 agaaaaacca agaaaagaca actggaacaa aaaatcagtt ccaaactagg aagtttaaag  
1500

ctgatggaac aggatacttg caatcttgaa gaggaagagc gaaaagcaag taccaaaatc  
1560

10 aaagaaataa atgttcaaaa agcgaaactt gttaccgaat taacaaacct aataaagatt  
1620

tgtacttctt tgcatatata aaaagtagat ttaattctcc aaaatactac agtgatctct  
1680

15 gagaagaaca aattagaatc agattalatg gccgcactct cacaactccg tcttacagag  
1740

caacatttca ttgaattgga tgaaaataga cagagattat tgcagaaatg caaggaaactt  
1800

20 atgaaaagag ctaggcaagt atgtaacctg ggtgcagagc agactcttcc tcaagaatac  
1860

cagacacaag taccaccat tccaaatgga cacaactccl cactcccat ggttttccaa  
1920

25 gaccttccaa acacallgga tgaaattgat gctttattaa ctgaagaaag atcaagagct  
1980

tcctgcttca cgggactgaa tcttacaatt gttcaggaat atacaaaaag agaagaagaa  
2040

30 atagaacagt taactgagga actaaaggga aagaaagttg aactagatca atacagggaa  
2100

aacatttcac aggtaaaaga aaggtggctt aatcclltaa aagagctggt agaaaaaatt  
2160

35 aatgaaaaat tcagcaattt ttttagttcc atgcagtgtg ctggtgaagt tgatctccat  
2220

acagaaaatg aggaagatta tgataaatat ggaattcgaa ttagagtcaa atttcgaagt  
2280

40 agtactcaac tgcatgaatt aactcctcat catcaaagtg gaggtgaaag aagtgtttct  
2340

accatgttat acttgatggc acttcaggag ctaaatagat glccattcag agtagttgat  
2400

45 gaaatcaatc agggaatgga cccaatcaat gaacggagag tgtttgaaat ggttgtaaat  
2460

actgcctgta aagaaaatac atctcaatac tttttcataa caccaaagct cctgcaaaat  
2520

50 cttecttatt ctgaaaagat gacagttttg tttgtctaca atggccctca tatgctggaa  
2580

ccaaacacat ggaatttaaa ggctttccaa aggcggcggc gccgtattac attcactcaa  
2640

55 ccttctlaat aaaagtaaag ayagggaaact tgggaatttt ttttgtaaa ttctgtttat  
2700

EP 1 355 149 A2

aagtatggct caactgaata aaaggagatt cactaaaacg aaaagcagtt atttttggaa  
 2760  
 5 acctgctttt aaatacaaat aggttgataa tggaaactat aatgaccttt ccaaaatagc  
 2820  
 agctggtagt aaaagttaag tcttcttcag tcttggttga acttgagttc ttggcactct  
 2880  
 10 gaccatgagt cattcagttc tcatgttaaa atgtacttaa tattacaaat caaaggatca  
 2940  
 gtggaagaag ggtaatcac aagaagttac ttatalggtt gccctgagct ttaattgcag  
 3000  
 15 agtaacttta attactttta gagcctaaag atgactctag agcctaagtc ctagtctctc  
 3060  
 ccaatgttat atttaatttt aaaaaattga tatgaaaatg tctaattgat agtaataatt  
 3120  
 20 tatgacagat ctagtcatth cttctcttta aaaaagatta ccttatctcc agtaggaaat  
 3180  
 ggaattttat gggcctttta aagaagttt tatgaaactt gatgctataa ttttattggt  
 3240  
 25 atttcaaggg gaaaaaagca ctgggggttca aaaatggtag cagaactgct ttgaaatgct  
 3300  
 gcaagggtgc cactagatga tgcaaaatac aaccaaaga ttgactgaga ataaaattag  
 3360  
 30 gtgacaaggg tttttaaaga ataacctttt aaagtgtggg ggcaggggtt gctttttttt  
 3420  
 attttattha aagtcaatta tattttacat cttacatttc taaaagcatt ttataattat  
 3480  
 35 ttttaqtaag atttttctta aaatttcata tactggtttc tacaatttat atttgaattt  
 3540  
 tctcagtytt atgtaaagag tgatggaaaa gcattgatth ctttaaaacc gtaatgtttt  
 3600  
 40 tagaacttha gcctataggg cttttcttac aatgttgatg taccatttat cttagaattt  
 3660  
 ctagtthaaa ctgttttctt tcaccgcaaa agaattaaat gggaaaatca tttgtttatc  
 3720  
 45 tctaagttat actaattagt agaaccaaac aaattatctt cttttaaaaa ataaatctta  
 3780  
 taggaaaaa gacagtccaa agtcatgtct ttgaacagtg gattggatct gtgccagtaa  
 3840  
 50 tgacaaaatt atttttttga cttgcttgcc tgaataaatt gaagaattgc tttcagtttg  
 3900  
 ggttttgtat attcttaagt agccattgaa atttatattc ttaactaggt caaaaaataa  
 3960  
 55 tgagccataa gtttatgtcc tctcacttag acattttctc tttaaaaagg tatttttctc  
 4020

tttataaaca ttttaaaaga gccttcctct cttaaactaa ctccagtgca tgaagtgga  
 4080  
 5 aaatatttta aaatgacatt ttactaata tgagcaagtc atgtaacat tgaagaactt  
 4140  
 ggtaacatat tagtaaatgg atattaccaa atgttttcat cgttaattac ttgcgttcc  
 4200  
 10 accaaaatat ctttactaaa atgtgcttgg ttagtattgt ttatlgcta aattagtlacc  
 4260  
 agtcatttta tttctgcaa atgagtatca atgtgaaaa gacacgtgaa gattaagcat  
 4320  
 15 gtttgaaaat aaaatggta attacatttc aatttacata ggccaacaac tgttccatac  
 4380  
 tttgtttgta aacatttaatt ttctctactg gacaaaatta atatttggtt ttacattgaa  
 4440  
 20 ttttgagctg tgaagaataa attatgtatc attttagcat attaaacagt agtaagtcta  
 4500  
 gacatagtc tcagccactt aaaacaaaag tttttttgtt tgtttgtttg ttgtttttt  
 4560  
 25 tgagatggag tctcactctg ttgccaggc tggagtgcag tggcgtgac tcggcttact  
 4620  
 gcaacctccg cctccgggt tcaagcgatt ctctgcctc agcctccca gtaactggga  
 4680  
 30 caacaggcgc gtcccaccac acccagctaa ttttttatac ttttagtaga gatggggttt  
 4740  
 cagcatattg gccaggctgg tctcgaactc ctgacctgt gatccaccg cctcggcctc  
 4800  
 35 ccaaagtgtt gggattatag gcgtgagccc ctgcaccgg ccaaaagttg atttttaatt  
 4860  
 acataaaaat cgtaaaaact tctagtaaaa acttgatttg gtgaatacag ttatatatta  
 4920  
 40 aaaccttaag gtgacaagca ttttctatgc ctaaatcttc attgggtttg ctggaaagag  
 4980  
 tctctgttaa aagattttcc atattcaaag taaaaggaaa gatttcttgc ttctaattg  
 5040  
 45 tcttttgag acatgcctat tttctttgag gtataaacct ttagatgtga aaaatgtaat  
 5100  
 ttcatctgc tattgtgtgt gcttgtgtgt gtgtaattga aaaaactggg aaatcctgct  
 5160  
 50 ttgttggtta taaatcaata tttttatatt c  
 5191  
 <210> 6  
 <211> 4755  
 55 <212> DNA  
 <213> human

EP 1 355 149 A2

<400> 6  
aagagatctt ccaggctctc agagccctgg gagggcgatt tccaggaaga ccacaatgcc  
60  
5 aacctctgga ggaggctgga gagagaaygc ctaggccaga gcctgtcagg caactttggc  
120  
aagaccaagt cagccttctc atctctccag aacattcctg agagtctgag aagacacagc  
180  
10 agcctggagc taggccgggg aaccaggag gggtaccccg ggggcaggcc cacctgtgca  
240  
gtcaacacca aggcagaaga ccctgggagg aaagccgctc ctgacctcg gagccatctg  
300  
15 gaccggcagg ttctctaccc gcggcccgag gggaggaccg gtgcctcggc ttctttcaac  
360  
agcacagacc caagtcccga agagccgcct gccccctcgc acccgcacac atccagtctg  
420  
20 ggccggaggg ggcccgcccc aggcagcgcc tcggctcttc agggctttca gtacgggaag  
480  
ccccactgct cgggtctgga gaaggtctcc aaatlcgagc agcgagagca agggagccag  
540  
25 agaccgagtg tgggcggctc tggttttggc cataactata ggccccacag gaccgtctca  
600  
acttccagta ctcttggaag tgacttcgag gagacaaaag cacacattcg ttctctgag  
660  
30 tcagctgaac ccctaggcaa cggggagcag cacttcaaaa acgggyagct gaagtggaa  
720  
gaggcttccc ggcagccctg cygllcagcag ctgagcgag gagcgtcggg cagcggccgt  
780  
35 ggccccaga ggccggacgc tcggctctc cgtagccaga gcaccttcca gctctccagc  
840  
gagccagaga gggagcccga gtggcgggac aggcccggt cggccgaatc gcccctgctg  
900  
40 gatgccccct tcagccgcgc ctaccggaac agcatcaagg acgcacagtc ccgtgtcttg  
960  
ggggccacct cctttcgacg tcgagacctg gagctggggg cycccggtggc gtccgaggtcc  
1020  
45 tggcggccac ggccttcctc ggcccacgtg gggctgcgga gcccagaggc gtcggcctcc  
1080  
gcctccccgc acacgccccg ggagcggcac agcgtgacct ctgctgaggg cgacctggcc  
1140  
50 aggccctgtc cccctgccgc ccggagaggt gctcgccggc gcctgactcc cgagcagaag  
1200  
aagcgtcct actcggagcc cgagaagatg aacgaggtgg ggtcglgga ggaggccgaa  
1260  
55 ccggcacccc tgggcccga gagaaatgg atgcgtttcc cggagagcag cgtggccgac

1320  
 5 cggcgccgtc tcttcgagcg cgatggcaag gectgtcca cgctcagcct gtcggggccc  
 1380  
 gagctgaagc agttccaqca gagcgccctg gcggactaca tccagcgcaa gaccggcaag  
 1440  
 10 cggcctacct ccgcgcggcg ctgcagcctc caggagcccg ggccactgcg tgagcgcgccc  
 1500  
 cagagtgcct acctccagcc cggccccgcg gcgctcgaag gctccggcct cgcctcgccc  
 1560  
 15 tccagcttga gctcactgcg ggagcccagc ctgcagcccc gcagggaggc caegctcctg  
 1620  
 ccggccacag ttgcagaaac ccagcaggct ccccgagatc gcagcagctc cttcgccggt  
 1680  
 20 ggccgcgccc tcggggaacg gcgacgcggg gacctgctta gcggagcaaa cgggtggaaca  
 1740  
 aggggcaccc agagagggga tgagaccccc agggagccat cctcctgggg ggccagggcc  
 1800  
 25 gggaagtcca tgcggccga ggacctgctg gaacgctcgg acgtccttgc gggccctgtc  
 1860  
 catgtgaggt ccaggtcatc tcccgccacc gcagacaagc gccaggatgt gcttttgggg  
 1920  
 30 caagacagtg gcttttgtct tgtgaaggat ccatgttatt tggctggtcc tggatctaqq  
 1980  
 tctactagtt gttcagaaag aggccaagaa gagatgctgc tgctcttcca ccatctcacc  
 2040  
 35 cctcgttggg gtggttcagg ctgcaaagcc attggtgatt cctccgttcc tagtgaatgt  
 2100  
 cctggaaccc tggaccatca gaggcaagcc agtaggacac cctgccccag gccaccactg  
 2160  
 40 gcaggaacgc aagggtcgtt cacagacacc agggctgcac cctgacccc aattggcacc  
 2220  
 cctctgcctt cagccattcc ctctggctac tgctcacagg acggtcagac agggcgacag  
 2280  
 45 cctctccgc cctacacccc tgccatgatg cacagaagca atggtcacac cctgaccag  
 2340  
 cctcccggtc caagaggctg tgaggcgat ggcccagagc atggggtaga agagggaaacg  
 2400  
 50 aggaagaggg tctcgctgcc tcagtgcca cctccttctc gagcaaagtg ggcccacgca  
 2460  
 gccagagagg acagccttcc tgaggaatcc tcagcccctg attttgaaa cctgaagcac  
 2520  
 55 tatcaaaaac agcagagtct tccaagtta tgcagcactt ctgacccaga cacacctctt  
 2580  
 ggggccccga gcactccagg gaggatctcc ctccgaatat ctgagtctgt cctgcgggac

2640

5 tccccgccac ctcatgagga ttatgaagac gaagtgtttg tgagggatcc gcaccccaag  
2700

gccaagtcga gcccacatt tgaacctctt cccccacccc cacctcctcc accgagtcag  
2760

10 gaaaccccg tglatagcat ggatgacttc cctccacctc ctccccacac tgtatgtgag  
2820

gcgcagctgg acagtgagga tcccagggg ccacgcccc gcttcaacaa actttctaaa  
2880

gtgacaattg caagggaaa gacatgcct ggtgcagccc atgtggtagg tagtcagaca  
2940

15 ctggcttcca gactccaaac lctatcaag ggttcagagg ctgagtcac accacctcc  
3000

ttcatgagcg ttcacgcccc acttgctggg tctcttggtg ggcagccagc acccatccag  
3060

20 actcaaagcc tcagccatga tccagtcagt ggaactcagg gtttagaaaa gaaagtcagt  
3120

cctgatcctc agaagagttc agaagacatc agaacagagg ctttgccaa ggaaattgtc  
3180

25 caccaagaca aatctctagc agacatcttg gatccagact ccaggctgaa gacaacaatg  
3240

gacctgatgg aaggtttgtt tccccgagat gtgaacttgc tgaaggaaaa cagtgtaaag  
3300

30 aggaaggcca tacagagaac tgtcagctct tcaggatgtg aaggcaagag gaatgaagac  
3360

aaggaagcag tgagcatgtt ggtaactgc cctgcctact acagtgtgtc tgctcccaag  
3420

35 gctgagctac tgaacaaaat caaagagatg ccagcagaag tgaatgagga agaggaacag  
3480

gcagatgtca atgaaaagaa ggctgagctc attggaagtc tcaccacaa gctggagacc  
3540

40 ctccaggagg cgaaggggag cctgctcacg gacatcaagc tcaacaacgc cctgggagaa  
3600

gaggtggagg ctctgatcag cgagctctgc aagcccaatg agtttgacaa gtataggatg  
3660

45 ttcatagggg atttggacaa ggtggtcaac ctgctgctct cctctcggg gcgtctagcc  
3720

cgtgttgaga atgtccctag cggccttggg gaagatgcca gtaatgaaga aaggagctct  
3780

50 ctttacgaga aaaggaagat cctggctggg cagcatgagg atgccggga gctgaaggag  
3840

aacctggatc gcagggagcg agtagtctg ggcattcttg ccaattacct ttcagaggag  
3900

55 cagctccagg actaccagca ctctgtgaaa atgaagtcca cgctcctcat tgagcaacgg

EP 1 355 149 A2

3960

5 aagctggatg acaagatcaa gctgggccag gagcaggtea agtgtctgct ggagagcctg  
4020

ccctcagatt tcattcccaa ggctggggcc ctggctctgc ccccaaacct cactgagtga  
4080

10 cccattcctg ctgggggctg tactttcagt ggtattttcc caacattaac ctctccactt  
4140

taacctcttc taaaataccc aacaaaaaga tcaactgttc tctcaacact atttaacttg  
4200

15 aaaaaatgtt cagtacaaac cactgtttga actatctggg ttattgggtg ttgttcctga  
4260

tgaaaggaaa aaaattctct ccaggaggaa gcctttttcc ttcttgccct tctgattga  
4320

20 tcttctgaga gctcgaatgc tgctggacac gtacccttt ctattattac tttgtagtag  
4380

aaagaaagt aatgaaactg agaactgatt ggagggtgtt tgatcattta gtttttaaca  
4440

25 ggctgaggca acatggatca gtgtgtgtcc ccttcaggaa tgtatccaca gtggccttcc  
4500

ttgtctgtgg gcagtgtatc ctgatggcag ggtacaagta ccattaatga agggctctga  
4560

30 acataaagcc ttaaaaagac acacactaag aaaactgtaa aaccttgaac attgttattt  
4620

atatttttta aaatgaaaa gatcactatg ttgtgtgtgc taaccactta tttgattctg  
4680

35 tttgtggtg gacalagatg attacgtttg agctttgtat tttgtgaaaa ccttaatgaa  
4740

atgaattcca aagat  
4755

40 <210> 7  
<211> 2045  
<212> DNA  
<213> human

45 <400> 7  
gaaacttgac cccggctcat cctgtctctg gctgtggccc ggcaaagcac tgaaaacccc  
60

tctggtctca gagacagtag gggcagtgcc actttctaca acctgccaac ccacacactg  
120

50 gagtaattct gaaaaaatt attcctaaac tctctaagtg tggacggaga atgagcaagc  
180

cccagaagta ttttacaacc agagtgggta atgaggaggg ggcttactgg aatcgtcata  
240

55 tctctgaata ttgaaaaca caactaaaa agtggacctt ctcaaaaaa aagggcagca  
300

EP 1 355 149 A2

aatgaccaag ggcgcccctt ctggccgtgc ttggcttgag taactgtctc tctttcccca  
360

5 ccccatcac agggctttca gtttggaaca ggaaaagcag ataaaaacag aacattccat  
420

atgtttcttt ctccatcggc caaaaacatt ttgacacaat gtttgtgaaa cacctttgga  
480

10 gaggtgcaact tctgaatgct gcctctgccg taaatcctgg ggcagggat cagcctcttc  
540

ccaggaacca tcgccttcta taaacgtga actcaagcag gcattttttt tttcttaccg  
600

15 aaaggctgct attgtgcaag ggcacataat gggctctgtg ctcttattgg cttccaaatg  
660

tgcatggcaa agagagagat gtgggcctag agcagalata ttcagcaagg tgacagcttc  
720

20 ccataacaat tctaacactt cttalcttat gtgagaataa aatatttaag ggttgaacct  
780

tatttlgcca aatgtatctt ttctgctttt gaattgggca gaagatttta gcaactatat  
840

25 tctacaaatg ttacttataa cacacacaca cacatctgaa atatatgccg aaaattgacg  
900

tctttgacct cagggagagc acctgtccag gctgcctaa aggaaatggc tccagtgggt  
960

30 ctaacaacc acatcctatc catggatagg tctagtata acactttaga gagaatgtca  
1020

gagcaggagg gaggcaagcc gcctcttctc ggccatcaac tgcagatgat gaaagagcgg  
1080

35 gattcaactt tgttttcttt tctgtggcc ccagtgaac ctctgccct cctgcacgt  
1140

ctgtgtcttc atttctaaaa tgggggtgat gctttcatat tgacctcacc ccatactacc  
1200

40 tcacagatgt gttgtgagga ttaataaaat tatgtctatg gtattttcag tttctggaga  
1260

aaaatactta tagacagttt aactattaca tagatatata agtgatctca gtttcttgtt  
1320

45 tgctgtgata ctaatgtgtt gttttaactt attccataaa atgacagttg tglcctagcc  
1380

acatcagaca gctatctaag ctctggacta cccctttgtg cagctgaatc actgcagggt  
1440

50 cgacctgcc tgggtgccaca gccatggttt ccatttctag atgaaaggat ggcctaggac  
1500

ataggtctca aagactcttg gatcagaatc aggagattag ggaaaacagg atggatacct  
1560

55 gagcactaac agcagtagac gtagacctct gtcctttacc atctgaggtc ttctggattc  
1620



EP 1 355 149 A2

tttgtgggt taattttgat ttgatgtcat ctgtttgcc ttcattctgc ttgcaagtgt  
 1680  
 5 gcatggttca atccclcaca tccaggaaat gaattttgca attyggccag atgctaattt  
 1740  
 gcacgttgat tcaccttctt tgcctttaag ccttpttttt cttttttttt ttttttgca  
 1800  
 10 aatgaatgta ccatttcaac tttgatttta atagtgetag ttgatattgg taataatgct  
 1860  
 aaccaagaga tcaatgccag atttttctct tggggtaagt tagctgaagt catttaaaga  
 1920  
 15 tggaaagggt ggaaaattct ttgatatttg atgtcattgt atccacattt gttgtaagac  
 1980  
 atattgcata ccaattataa ttatatcaat taaagttgat aaaagcttca aaaaaaaaaa  
 2040  
 20 aaaaa  
 2045  
 <210> 8  
 <211> 2096  
 <212> DNA  
 25 <213> human  
 <400> 8  
 atggagaacg agcctgtagc ccttgaggaa actcagaaga cagatcctgc tatggaacca  
 60  
 30 cggttcaaag tggtaggattg ggacaaggac ctggtggact ggcgaaagcc tclctgtgg  
 120  
 caggtgggcc acttgggaga gaagtacgat gaglgggttc accagccggt gaccaggccc  
 180  
 35 atccgcctct tccactcaga cctcattgag ggcctctcta agactgtctg gtacagtgtc  
 240  
 cccatcatct gggtagccct ggtgtgtgat ctacagctgt cctactaccg aacctttgcc  
 300  
 40 cagggaacg tccgactctt cagtcattt acaacagagt acacgggtggc agtgcccaag  
 360  
 tccatgttcc ccgggtctt catgtgggg acatttctct ggagcctcat cgagtacctc  
 420  
 45 atccacgct tctgtttcca catgaagccc cccagcgaca gctattacct catcatgtctg  
 480  
 cacttctgta tgcacggcca gcaccacaag gcacccttcg acggctcccg cctggcttc  
 540  
 50 cccctgtgc cagcctccct ggtgatggc gtcttctact tgtgatgca gctcatctg  
 600  
 cccgaggcag tagggggcac tgtgtttgcg gggggcctcc tgggtacgt cctctatgac  
 660  
 55 atgaccatt actacctgca ctttggctcg ccgcacaagg gctcctacct gtacagcctg  
 720

aaggcccacc acgtcaagca ccactttgca catcagaagt caggatttgg tatcagcact  
780

5 aaattgtggg attactgttt ccacaccctc actccagaga aacccacact gaagacgcag  
840

tgacaactcc caccacctcc gtctgtccct cagcccggcc ctggccctt cccgaccccc  
900

10 acccgccatt cagaccccat taagaaggtt ggcttgcca gccaggatgg gctgtgtccg  
960

gccctgcagc ctagtgaag gtgctgaggg ggccctgagc caggaccgcc ctctgaccc  
1020

15 ctggtaggag ggtcacatcc acttggtgca ggtggccctt ggtgaccac ttcttctgg  
1080

agcgtccctg cctagagctc agcccacagg actgcttcag gccgtggcca caggtagcag  
1140

20 ccgcaagggg aaatgaagaa aactgagccc tcgtggccac ctgtgtcacc cttgtgcctt  
1200

agcctcatgg gctgcctagg agctgcctgc acggcacagc tcgctttcac agtcagaagt  
1260

25 gggctctgtg gatctgtggt ccctgtctc cctgctgtcc ctctctggga ggctttggtg  
1320

gctctgaggt ggacaaagag ctctcgcaag aagagacagc gtgatgcctc ccacagtcca  
1380

30 cccagacccc tggggcagcc cctctggccc tgccagctgc ctgcgtcgtt gggcccaggg  
1440

tggttgagcag gagtcccagc tgcttgcctt aggcctggc agcttttctt gccgtccctc  
1500

35 ccctgcctcc agaattcacag cccttctccc caaggagggc tgaggaggct tctccaccag  
1560

tggcagcccc acccgtccc tgccattct tgccctccac cccgtcagg cccctactcg  
1620

40 ggcgtcccca gaaggagcca cctctcagtg cctcacctcc ccctgcctcc cagcctccgc  
1680

agatgaggtt cctgcccctt cctcctcgta accaaaaacc tcaactgtcc caggacggtc  
1740

45 ttatttataa accagataca tgttcttagt ctggtcccag accaaggagc tggtcagacg  
1800

gccccttcta atctacatg ttgagcttat gtaaaaaatg ttgttctctc ctgtttttgg  
1860

50 ttcccttctt acccacaacc cattactact tgaaacttaa aaaactcgcc aagtgtaaag  
1920

gctaaagaga agcagtttga cggacctgt gatttgtact gttgtgtcg gagctattta  
1980

55 aagattttgg aataaatata caaaactacg gttgtgaaat aaaaacttaa attgtatatt  
2040

ttgaaaaata aaacactgaa aagaaaccaa caaaaaaaaa aaaaaaaaaa aaaaaa  
 2096

5

<210> 9  
 <211> 5640  
 <212> DNA  
 <213> human

10

<400> 9  
 ggaaacgcag aaacagaga gaggcattct gagtcattctg actggatgaa gactgttcca  
 60

agttacaacc aaacaaatag ctccatggac tttagaaatt atatgatgag agatgagact  
 120

15

ctggaaccac tgcccaaaaa ctgggaaatg gcctacactg acacagggat gatctacttc  
 180

attgaccaca ataccaagac aaccacctgg ttggatcctc gtctttgtaa gaaagccaaa  
 240

20

gccccgaag actgtgaaga tggagagctt ccttatggct gggagaaaat agaggacct  
 300

cagtatggga catactatgt tgatcacctt aaccagaaaa ccagtttga aaatccagt  
 360

25

gaggaagcca aaaggaaaa gcagtttaga caggttgaaa ttgggtcttc aaaaccagat  
 420

atgaaaaat cacacttcac aagagatcca tcccagctta aaggtgtcct tgttcgagca  
 480

30

tcaactgaaa aaagcacaat gggatttggc ttactatta ttggtggaga tagacctgat  
 540

gatttcctac aagtgaaaaa tgtgctgaaa gatggtcccg cagctcagga tgggaaaatt  
 600

35

gcaccaggcg atgttattgt agacatcaat ggcaactgtg tcctcggta cactcatgca  
 660

gatgttgccc agatgtttca attggtacct gtcaatcagt atgtaaacct cactttatgt  
 720

40

cgtggttacc cacttcctga tgacagtga gatcctgttg tggacattgt tgctgtacc  
 780

cctgtcatca atggacagtc attaaccaag ggagagactt gcatgaatcc tcaggatttt  
 840

45

aagccaggag caatggttct ggagcagaat ggaaaatcgg gacacacttt gactggtgat  
 900

ggtctcaatg gaccatcaga tgcaagtga cagagagtat ccattggcacc gtcaggcagc  
 960

50

tcccagcctg aactagtga tatccctttg attaagggcc ctaaaagggt tgggtttgca  
 1020

attgctgaca gccctactgg acagaaggty aaaatgatac tggatagtca gtggtgtcaa  
 1080

55

ggcccttcaga aaggagatat aattaaggaa atataccatc aaaatgtgca gaatttaaca

1140  
catctccaag tggtagaggt gctaaagcag tttccagtag gtgctgatgt accattgctt  
1200  
5 atcttaagag gaqgtcctcc ttaccaacc aaaactgcc aatgaaaac agataaaaaag  
1260  
gaaaatgcag gaagtttggg ggccataaat gaggctattc ctcagcctat gccttttcca  
1320  
10 ccgagcatta tcaggtcagg atccccaaaa ttggatcctt ctgaggctta cctgaaatct  
1380  
aagactttat atgaagataa accaccaaac accaaagatt tggatgtttt tcttcgaaaa  
1440  
15 caagagtcag ggtttggtt cagggtgcta ggaggagatg gacctgacca gtctatatat  
1500  
attggggcta ttattccctt gggagcagct gagaaagatg gtcggctccg cgcagctgat  
1560  
20 gaactaatgt gcattgatgg aattcctgtt aaagggaaat cacacaaaca agtcttggac  
1620  
ctcatgacaa ctgctgctcg aaatggccat gtgttactaa ctgtcagacg gaagatcttc  
1680  
25 tatggagaaa aacaaccgga ggacgacagc tctcaggcct tcatttcaac acagaatgga  
1740  
tctccccgcc tgaaccgggc agaggtccca gccaggcctg caccocagga gccctatgat  
1800  
30 gttgtcttgc aacgaaaaga aatgaagga ttgggctttg tcctcctcac ctccaaaaac  
1860  
aaaccacctc caggagtat tctcatalaa attggccgag tcatagaagg aagtcgggct  
1920  
35 gaccgctglg gaaaactgaa agttggagat catatctctg cagtgaatgg gcagtcatt  
1980  
gttgaactgt ctcatgataa cattgttcag ctgatcaaag atgctggtgt caccgtcaca  
2040  
40 ctaacggtea ttgctgaaga agagcatcat ggtccaccat caggaacaaa ctcagccagg  
2100  
caaagcccag ccctgcagca caggcccatg ggacagtcac aggccaacca catacctggg  
2160  
45 gacagaagtg ccctagaagg tgaaattgga aaagatgtct ccacttctta cagacattct  
2220  
tggtcagacc acaagcactt tgcacagcct gacaccgag taatttcagt tgtaggcagt  
2280  
50 cggcacaatc agaaccttg ttgttatcca gtagagctgg agagaggccc ccggggcttt  
2340  
ggattcagcc tccgaggggg gaaggagtac aacatggggc tglcatcctt tcgtcttgct  
2400  
55 gaagatggtc ctgccatcaa agatggcaga attcatgttg gtgaccagat tgttgaatc

EP 1 355 149 A2

2460  
aatggggaac ctacacaagg aatcacacat actcgagcaa ttgagctcat tcaggctggt  
5 2520  
ggaaataaag ttcttcttct tttgaaggca ggaactggct tgatacctga ccatgggttg  
2580  
gctccttccg gtctgtgctc ctacgtgaaa cccgagcaac attaaggctt tcagggcttt  
10 2640  
tcttggtctt tccttaaaaa gacttggtga ttgggatatt aataatcctt cgtcttcaaa  
2700  
tgtgatttat gatgaacagt caccattacc cccatcttca cattttgctt ccatatttga  
15 2760  
agagtctcac gtgccagtaa ttgaagaatc ttlgagagtt cagatatgtg aaaaggcaga  
2820  
agaattaaag gacattgtgc ctgaaaagaa aagcacttta aatgaaaatc agcctgagat  
20 2880  
aaagcatcag tctcttctcc agaaaaatgt gagtaagagg gatccacca gcagtcatgg  
2940  
gcacagtaac aagaaaaatc tattaaggt agaaaatggt gttacacgaa gagglagatc  
25 3000  
ggttagctcc aaaaagccag ccagtcaaca ttcagaggaa catttgata agattcctag  
3060  
tcctctaaaa aataacccca aaagaagacc cagagatcaa tccctcagcc ccagcaaagg  
30 3120  
ggaaaataaa agttgtcagg tcagcaccag ggcaggctct ggacaagatc agtcagaaaa  
3180  
aagcagaggt cggtcggcca gcccaaaaaa gcagcaaaaa attgaaggaa gcaaagctcc  
35 3240  
atcaaatgct gaggccaaal tattagaggg taagagtcga agaatagcag gctatacggg  
3300  
cagtaatgct gagcagatcc cagatgggaa ggaaaaatca gacgtcatca ggaaagatgc  
40 3360  
aaagcagaat cagttgaaa aaagcagaac aaggtctcca gagaaaaaaa tcaaaagaat  
3420  
ggttgagaaa tctcttccat ccaaaatgac taataagact acaagtaaag aagtatctga  
45 3480  
aatgaaaaa ggaaagaaay taaccacagg agaaacaagt tctagtaacg ataaaatagg  
3540  
agaaaatgct cagctatcag aaaagaggct gaagcaagaa cctgaagaga aggtagtctc  
50 3600  
aaacaaaaca gaagatcaca aagggaaga actagaggca gctgacaaaa acaaagagac  
3660  
tggaaggctc aaaccgaaa gcagttctcc agttaagaaa aactgataa ctccagggcc  
55 3720  
ctggaagggt ccaagtggaa ataaagtcac aggcactall ggtatggctg agaaacggca

EP 1 355 149 A2

3780

5 gtaaccttta gtataaaaca aagaaaaaca agttgtaatc ttttcttaca gcagcatttt  
3840

tccagaaaaa gccttttttt ttttttcaga tattctgaaa cagataagta catgttaatg  
3900

10 tgagcctcaa gtlacctagg ctgcatgaag ggccttttagg attgctaaga accaactgtc  
3960

ccccggccg gctgcctcc ctcgctctca ggaaggagct gcacccacat gctcatctga  
4020

15 cccgccctgc tcaggctgcc cagctcgtct tcatgagtgt ctgaacaaat gacatatgtt  
4080

gatattaaca atgtggtcac aactcacttt gtatttgtgc caagtatatct actgtatcat  
4140

20 gtctgttttt atcctttttt ttcagctggt tccacagtaa tgaaaaagtt aggtttggct  
4200

tggaagtga tgatctcaat agcatgttgc atgtttacag agagaaatat gtgagtcctt  
4260

25 gcagaagaag agactgttaa ctcacgttta aaqatggccg ttgtctcttc taacagctac  
4320

tgatgatgic ccactttaaa aataaaaccc ccaaacatca ctactttaag gaaaaaaaaa  
4380

30 atgtagtcca atattgatgc tttcttatgg ctttttatct taatttggct ggataagttg  
4440

tttcaataa ctgttaaga tattacttac aattgaatgt ttgaaataag aaagtacttt  
4500

35 aagcaataga gttcatctcc tgctgtgtta tccaacctcg atglatactt acagcatctc  
4560

aggtcacctt ttttatttca gttatttaat tatgaaacca taaagaagca tgtggaaata  
4620

40 gtgtttattg ctctttgaag aaaaaccacc aactatttct ggatattttg gctgtaccta  
4680

ctactaaagt cattagtctt taatacataa tacatatattg aaaagtaaac atattatata  
4740

45 gattatgtga gggacttaat catgaaacca gtttcacagt ccaagtacca actcttctgg  
4800

tagcaggtgc acaagcttgg gtgtttaaaa acaacctgtg tagggatgc ccagcaaatg  
4860

50 aggacaaatg tgtagacagt acttactgga tcttatttaa cttttagcta cattaactaa  
4920

ctttcttatt taaaaacaag aaagggagac taaacatctg cttaacttgt acacatttcc  
4980

55 agaattcttt taaaagtct agttaaagat gtttcttaga agttggagac tgttaacaac  
5040

ttccataaaa tagatccagg tttttcagtt cctgaagca gcattcagta gcctctatat

# EP 1 355 149 A2

5100

aaataaaggc accttctgag aataaaacta ttttatggag tgtgtgaaca cacttgttct  
5160

5 gtcacctggg ttcattctgt tgtgaagcac attaggtcca ggcccttccc tctgggagtc  
5220

tgactgtgaa actctllaac ccaacaactc aattagcccc ttagataag acatgcttcc  
5280

10 cagagtgaga tttttgaaat ccccttttca tccagaacta tatttaccba cctattgtaa  
5340

ctattcaaat agagcaaaat taggaggctt gataaatact aagaatttag taccacagaa  
5400

15 attatttatt attttccctg tagtcacaaa ttagtgataa cgaatcctat ttttgtaa  
5460

tgtagacataa ctttgatgtc atatgtgtc ctatgtggtt cttcctaagt aaactctgta  
5520

20 ctgattatat actgacttag caatgtggcc ttggaatgct gagcaaaatg tggatgtact  
5580

ggttgtaaat gtttatatat tgtacagtac ctttatatat acacttgagg ttctgattag  
5640

25

<210> 10  
<211> 457  
<212> DNA  
<213> human

30

<220>  
<221> misc feature  
<222> (242)..(242)  
<223> any kind of base

35

<220>  
<221> misc feature  
<222> (369)..(369)  
<223> any kind of base

40

<220>  
<221> misc feature  
<222> (394)..(394)  
<223> any kind of base

45

<220>  
<221> misc feature  
<222> (406)..(406)  
<223> any kind of base

50

<220>  
<221> misc feature  
<222> (457)..(457)  
<223> any kind of base

55

<400> 10

# EP 1 355 149 A2

tcagtcactc ttccaccctg ccaaagcttc actgtccctac tgattgaatt gtatgtgaga  
 60  
 5 aataaaatgt catcatatta agccactggg atttgtatgt ttatctgtta tagcagcaag  
 120  
 tcttaattta cctaatacac acattgtgac agatgttctt aatgtccac cccatattgt  
 180  
 10 tacatgtcca gctttgagga tccctggcat gtgggggtag gagtttctgg gcatgctgga  
 240  
 tncaattccc acttttaagg catctgtggc ctctgtggcc tctgtggcct tcaactgttat  
 300  
 15 ggaagggtt tatctggggc accataggaa actttaccat ggcacagtgg acaacctagg  
 360  
 agggggtgng gaggaggggc ctccaggccc aacngggggg accagngtcc gtggggtag  
 420  
 20 ggtggttttg ggggttttcc ctcttaeccg tgggggn  
 457  
 <210> 11  
 <211> 1493  
 <212> DNA  
 25 <213> human  
 <400> 11  
 aatagggttg gcggtgcag cgggcggcaa acagcccgcc cggcaccacc atgctcgccc  
 60  
 30 tggaggctgc acagctcgac gggccacact tcagctgtct gtaccagat ggcgtcttct  
 120  
 atgacctgga cagctgcaag cattccagct accctgattc agagggggct cctgactccc  
 180  
 35 tgtgggactg gactgtggcc ccacctgtcc cagccacccc ctatgaagcc ttcgaccgg  
 240  
 cagcagccgc ttttagccac cccaggtctg cccagctctg ctacgaacco cccacctaca  
 300  
 40 gcctgcagg gaacctcgaa ctggcccccga gcctggaggc cccggggcct ggctccccg  
 360  
 cataccccac ggagaacttc gctagccaga ccctggttcc cccggcatat gccccgtacc  
 420  
 45 ccagccctgt gctatcagag gaggaagact taccgttga cagccctgcc ctggaggtct  
 480  
 cggacagcga gtcggatgag gccctegtgg ctggccccga ggggaaggga tccgaggcag  
 540  
 50 ggactcgcaa gaagctgcgc ctgtaccagt tctgtctggg gctactgacg cgcggggaca  
 600  
 tgcgtgagtg cgtgtggtgg gtggagccag gcgccggcgt cttccagttc tcctccaagc  
 660  
 55 acaaggaaact cctggcgcgc cgctggggcc agcagaaggg gaaccgcaag cgcattgacct  
 720



accagaagct ggcgcgcgcc ctccgaaact acgccaagac cggcgagatc cgcaaggcca  
780

5 agcgcaagct cacctaccag ttcgacagcg cgtctgtgcc tgcagtcgcc cgggcctgag  
840

cacacccgag gctccacact gcggagccgc tgggggacct cacgtcccag ccaggatccc  
900

10 cctggaagaa aaagggcgtc cccacactct aggtgatagg acttacgcat ccccaccttt  
960

tggggtaagg ggagtgtgc cctgccataa tccccaagcc cagcccgggc ctgtctggga  
1020

15 ttccccactt gtgcctgggg tccctctggg atttctttgt catgtacaga ctccctggga  
1080

tctcatgtt ttgggtgaca ggacctatgg accactatac tcggggaggg agggtagcag  
1140

20 tgcttcaga gtcccaagag cttctctggg atttctttgt gatctctgat tcccagtgga  
1200

ggcctgggac ctttttaaga tcgctgtgtg tctgtaaacc ctgaatctca tctgggggtg  
1260

25 gggccctgct ggcaaccctg agccctgtcc aaggttcctt cttgtcagat ctgagatttc  
1320

ctagttatgt ctggggccct ctgggagctg ttatcatctc agatctcttc gccatctat  
1380

30 ggctgtgttg tcacatctgt cccctcattt ttgagatccc ccaattctct ggaactattc  
1440

tgctgccctt ttttatgtgt ctggagtcc ccaatcacat ctagggtcc tcc  
1493

35

<210> 12  
<211> 2292  
<212> DNA  
<213> human

40

<400> 12  
ccatgggttc cccttcagcc tgtccataca gagtgtgcat tccctggcag gggctcctgc  
60

tcacagcctc gcttttaacc ttctggaacc tgccaaacag tgcccagacc aatattgatg  
120

45

tcgtgccgtt caatgtcgca gaagggaagg aggtccttct agtagtccat aatgagtcct  
180

agaatcttta tggctacaac tggtaaaaag gggaaagggg gcatgccaac tatcgaatta  
240

50

taggatatgt aaaaaatata agtcaagaaa atgccccagg gcccgcacac aacggtcgag  
300

agacaatata cccaatgga accctgctga tccagaacgt taccacaat gacgcaggat  
360

55

tctataccct acacgttata aaagaaaatc ttgtgaatga agaagtaacc agacaattct

420  
acgtattctc ggagccaccc aagccctcca tcaccagcaa caacttcaat ccggtggaga  
480  
5 acaaagatat tgtggtttta acctgtcaac ctgagactca gaacacaacc tacctgtggt  
540  
gggtaaacaa tcagagcctc ctggtcagtc ccaggctgct gctctccact gacaacagga  
600  
10 ccctcgttct actcagcgcc acaagaatg acataggacc ctatgaatgt gaaatacaga  
660  
acccagtggg tgccagccgc agtgacccag tcaccctgaa tgcctcgctat gagtcagtac  
720  
15 aagcaagttc acctgacctc tcagctggga ccgctgtcag catcatgatt ggagtactgg  
780  
ctgggatggc tctgatatag cagccttggg gtagtttctg catttcggga agagtgtttt  
840  
20 tattatccac ctgcagactg gactggatc ttctagctcc ttcaatccca ttttctcctg  
900  
tggtatcact aagtataaga cctgctctct tcctgaagac ctataagctg gaggtggaca  
960  
25 actcaatgta aatttcaagg aaaaaccctc atgcctgaga tgtggggcac tcagagctaa  
1020  
ccaaaatggt caacaccata actagagaca ctcaaattgc caaccaggac aagaagttga  
1080  
30 tgacttcatt ctgtggacag tttttcccaa gatgtcccaa gcctcatcgt gacgaggctc  
1140  
ttatccact ccatttttcc ctgctcatgc ctgcctcttt aatttggtaa gataatgctg  
1200  
35 taactagaat ttcaaatca gcgccttggt caggcaattt gacagagtgt tggatgtgtc  
1260  
atgtcatcat glcaaaccca aatatttgac ctaagggatc ctttattctg ccagtggtc  
1320  
40 aactttaaca acatccctaa tacaactgtt tattcaaatg cacggtggtc cctgttagag  
1380  
ttagacctct agactcacct gttctcacgc cctgttttaa ttaaccag ctatgggatg  
1440  
45 ccagataaca gaattgctgc ctacgagctg aacagggagg agtttgtgca gttgctgaca  
1500  
cttcttggtg cacataaata aatacagtg gtactataga gactcagttg caaaaattaa  
1560  
50 caaatalgct gcttgattaa aatgggtagg ctctcatgt ggctcattct ttaatctatt  
1620  
ctcttttatt tggtttggt catggggtct ctgcctatgg atcatacttc aaactcttgg  
1680  
55 tgtgatcctc ctgattgtca caatattagt taccctgggt tgctgtattc tctaaaacct

# EP 1 355 149 A2

1740

5 ttaaattgttt gcatgcagcc attcgtcaaa tgtcaaatat tctctctttg gctggaatga  
1800

caaaaaactca aataaatgta tgattaggag gacatcataa cctatgaatg atggaagtcc  
1860

10 aaaaatgatgg taactgacag tagtgtaaat gccttatggt tagtcaaact ctcatttagg  
1920

tgacagcctg gtgactccag aatggagcca gtcattgctaa atgccatata ctcacactga  
1980

15 aacatgagga agcaggtaga tcccagaaca gacaaaactt tcctaaaaac atgagagtcc  
2040

aggctgtctg agtcagcaca gtaagaaagt cctttctgct ttaactctta gaaaaaagta  
2100

20 atatgaagta ttctgaaatt aaccaatcag ttattttaaa tcaatttatt tatattcttc  
2160

tgttcctgga ttcccatttt acaaaaccca ctgttctact gttgtattgc ccagtaggag  
2220

ctatcactat atlllgcaga atggaaactg ccctgactct tgaatcacia ataaaagcca  
2280

25 attgtatctg tt  
2292

30 <210> 13  
<211> 519  
<212> DNA  
<213> human

35 <220>  
<221> misc\_feature  
<222> (212)..(212)  
<223> any kind of base

40 <220>  
<221> misc\_feature  
<222> (451)..(451)  
<223> any kind of base

45 <400> 13  
gaaacaacaa cagtgtaatc tttaacaggg atgttaaagg taagaagtca ggaagataaa  
60

ccaaatgat tgagtatgat aaagaatttt gcatggcgat taaaatagaa aacctataaa  
120

50 ttagaataaaa gcaggctctgg acttagcaaa gaaacaatat agtttgagga aggcattgaa  
180

taagttcttt tcatgttcac tgctggtcac ancataacag agagtgatgt ggagagcttt  
240

55 ggggaaggttt cacgttgagt tacatcagtg gtcaacaatg gagcaacaag actccgtaga  
300

# EP 1 355 149 A2

ggatgccacc ctgggagaat tgcaagggaaggaggctga agcacaactg gtaatagcct  
 360  
 5 tcagatattt aatggatatg caaataaagc tctgattaat tgtattttca cttattatat  
 420  
 atcatctttg gacctttcta aaagtgggac nctagaaaag atatactgaa actccaaaag  
 480  
 10 aatacttcag ctcgagttga atggattcaa gatgttgtt  
 519  
 <210> 14  
 <211> 5294  
 <212> DNA  
 15 <213> human  
 <400> 14  
 ggctcgcac cccatagtc tgggttacag tgaaggtag ccccgcgctc tgctctggag  
 60  
 20 aggcagggg ggatagggaa cgtctcaggt ggcgcccga gtcatggtgg tgttcgttgg  
 120  
 ccgcgcctc ccggcgctcc tagggctgtt taagaagaag ggctctgcca aggctgagaa  
 180  
 25 tgacaaacat ctaagtgtag ggcctggcca ggggccaggg tctgcagtgg atgagacca  
 240  
 ggacaacgtc ttctttccca gtgggcgacc ccccccactg gaagagctgc aactcaggc  
 300  
 30 ccaggagggg ctccgctccc tacaacacca agagaaacag aaactgaaca agggtggtg  
 360  
 ggaccatga gacaccaga gtatccagtc ctcccgagc gggccggatg aagacaacat  
 420  
 35 ctctcttcgc agtcagacca catctacgt ggctgagagc tccacagcag aggcgcgct  
 480  
 ctccatccgc tcggagatga tccagcgcaa aggtccacc ttccgacccc atgactcatt  
 540  
 40 tcccaaatct gaaagtcag ggcgcytcy gcgggagcgg cggagcactg tgctgggact  
 600  
 ccgcagcat gtgcagaagg agcttggcct gaggaatgag cgtgaggcac caggcacgcc  
 660  
 45 ccgggtcct ggtgcacggg atgccgtacg catccccaca gtggacggcc gccccgagg  
 720  
 cacctcaggg atgggggccc ggtgtccct gcaggcgctg gaggcggagg cagaggctgg  
 780  
 50 cgctgagaca gaggccatgc tgcagcgcca cattgaccgt gtctaccggg atgacacctt  
 840  
 tgttggccg tccacgggta cccgggcccc accattgacc cggcccatgt ccctagcagt  
 900  
 55 gcctggattg acaggagggg cagggcctgc agagcccctg agcccgcca tgtecatctc  
 960

cccccaggcc acctaactgt cgaagttgat tccacatgct gtgctgccgc ctacagtggg  
 1020  
 5 cgtggtggcc ctaggccgct gcagcctgcg cacactaagc cgctgcagcc tgcactcggc  
 1080  
 cagcccagcc tcagtcgct cgctggggcg cttctcctcc gtctccagcc cacagccccg  
 1140  
 10 cagccgccac ccatacctct ccagtgcac ctggagccac tctcaatct ccgacaccat  
 1200  
 tgtgtctgac ggttcacccc tctcctctaa gggtagctct gagggccagc cggagagctc  
 1260  
 15 tacggctagc aatagcgtgg taccctctcc ccaggaggc agtgggaggg gctctccag  
 1320  
 tgggggcagc actgctgagg cctcagacac actcagcatt cggagcagtg ggcagttgtc  
 1380  
 20 tggccggagt gtgtccctgc gtaagctgaa gcggcctcca cccctcccc gccggacca  
 1440  
 ctccctccat cagcggggct tagcagtgc tgatgggcca ttagggttgc cccctaagcc  
 1500  
 25 tgagcgtaag cagcagcccc agctgcctcg gccaccacc actggtggct cagaaggggc  
 1560  
 gggggcagca ccctgtccac ccaaccagc caacagctgg gtacctggt tgtctccygg  
 1620  
 30 tggttcccg cgccccccac ggccccaga acggacactt tcgccctcca gtggatactc  
 1680  
 gagccaaagt ggtactccca cctccctcc caagggcctg gcaggctccc ctgcttccc  
 1740  
 35 aggcaaggcc cagcccccta aaccagagcg tgacacgtct ctctgctccc ctggggcctc  
 1800  
 cgtctcctct tccctcacgt ctttatgttc ctctcctct gaccagccc cctcagaccg  
 1860  
 40 ctctgggcca cagatattga cccctctggg tgacaggttt gtcatactc ctcaccccaa  
 1920  
 ggtgcctgcc cctttctccc cactccctc caagcccagg agccctaacc cagctgccc  
 1980  
 45 tgctctagcc gccctgctg tggttcctgg gcctgtttct accactgacg ccagtccca  
 2040  
 gtccctccc actcccaga caaccttgac tccactgcag gagtctctg tcattccaa  
 2100  
 50 agaccagtca cccccactt cccaccccc atcttatcat ccacccccc caccactaa  
 2160  
 gaagccagag gtggttgagg aggcaccatc tgcctcagag actgctgagg agccctcca  
 2220  
 55 agatcccaac tggccccctc cccaccccc tgccctgag gagcaggacc tgtccatggc  
 2280

tgacttcccc ccaccagagg aggetttttt ctctgtggcc agccctgagc ctgcaggccc  
 2340  
 5 ttcagggtcc ccagagcttg tcagctcccc ggctgcttgc tcctctcag ctactgcttt  
 2400  
 gcagattcag cccccgggta gcccagaccc tcctccagct ccgccagccc cagctcctgc  
 2460  
 10 tagttccgcc ccagggcattg tggccaagct ccctcagaag gaaccggtgg gctgtagcaa  
 2520  
 ggggtggtgg cctcccaggg aggacgtagg tgcgccctg gtcacgccct cgtcctgca  
 2580  
 15 gatggtgcgg ctgcgtccg tgggtgctcc aggaggggct cccaccccag cactggggcc  
 2640  
 atcgccccc cagaaaccac tgcgaagggc cctgtcaggg cgggccagcc cagtgcctgc  
 2700  
 20 cccctctca gggctccatg ctgcgtccg actcaaggcc tgcagcctgy ccgccagtga  
 2760  
 aggcctctca agtgtcagc ccaacggacc gcctgaggca gaggcacggc ctccccagtc  
 2820  
 25 ccctgcctca acggccagtt tcattctctc caagggtctt aggaagctgc agctggagcg  
 2880  
 gcccggtgcc cctgagaccc aggtgacct ccagcggaat ctggtggcag aactccggag  
 2940  
 30 catctcagag cagcggccac ccaggcccc aaagaagtca cctaaggctc cccacctgt  
 3000  
 ggcccgaag ccgtctgtgg gagtcccccc accgcctcc ccagttacc ctgcagctga  
 3060  
 35 gcccttact gctcctcca ccaatgggct ccctcacacc caggacagga ctaagaggga  
 3120  
 gctggcggag aatggagggt tcctgcagct ggtgggcca gaggagaaga tgggcctccc  
 3180  
 40 gggctcagac tcacagaaag agctggcctg accaccaggc acctactgg cactgctgac  
 3240  
 ccatcccaga aacacaatct cagggaaccc agcagctcca aggacgagag gatacagcag  
 3300  
 45 acacaacctc atagagaggg cgctgcagc cttaacctcc acggccttcg atacttatgc  
 3360  
 aagcctggtg ttgctcctgt cctcagagtc atcctgcgct catgcctttt cccgaatggg  
 3420  
 50 ttcacctctg gcagttgccg cttcagcttt ggccctagcc tcattctgaa gtggylagct  
 3480  
 ggccgggagag ggtggctgcg cccctgctg gccctgaggc tgcagagttg ggagcaggac  
 3540  
 55 acctcaactg agtttcattt ttttctatgt ccaaaccatg cacatactat agtcagaat  
 3600

caaagcactt ttgaaaagtg gctgcatggc catctctcag ggcccaggaa gttgcattcc  
3660

5 aagggcctgt ttacatygca gcagaatcca tccccggcag tcagcccata gcttgggacc  
3720

agtctygtcc ctccctgcca gtccagttta ctccctcttg ttcttgaagg tggccaagtc  
3780

10 attgtgttcc cacaggcttc tctaggctgg gggcaggtgt ggggctgtgg aattccaaag  
3840

cacaaaaggt gcagagggga ttggccttcc tgtgcctcaa ctaccaacc accctctcgc  
3900

15 cttccagttc tgccaggtgc tccatgctgg ggacaagtag gagactgcc aagggccaaag  
3960

aaatgggtga gcagtagagt catctcgggg cacttggcag tgtcaagcac ctgccccttg  
4020

20 cctccttgac cacactgggg tgggtgggcc cccagcactt cagaggcagg agcctttggg  
4080

ctgagcaagc actgaggagg tggatggaag ggagcatctg gaggggggga gcttccttga  
4140

25 gcagtgggcc caggcctggc cctccacact tcattctctg acctttctct ctctcattt  
4200

cggtgcatgt cctttctgca gctgccttcc agcacagggt gttccactgg gggcagctaa  
4260

30 cgctgagtga caaggatggg aagccacagg tgcattttac tcaagtcttc tctagtcaat  
4320

gaggggcacc cagtgcctct agggcaggct ggggtggtgt cccctaggta tcagcctctc  
4380

35 ttactgtact ctccgggaat gttaaccttt ctattttcag cctgtgccac ctgtctagtc  
4440

aagctggctt ccccatggc cctgtgggt ccacagcagc gtggtgccc cccagggcc a  
4500

40 ccgcttcttt cttgatctc tttccttaac agtgacttgg gcttgagtct ggcaaggaac  
4560

cttgctttta gcttcaccac caaggagaga ggttgacatg acctccccgc cccctacca  
4620

45 aggctgggaa cagaggggat gtggtgagag ccaggttcct ctggccctct ccagggtgtt  
4680

ttccactagt cactactgtc ttctccttgt agctaataa tcaatattct tcccttgctt  
4740

50 gtgggcagtg gagagtgtg ctgggtgtac gctgcacctg cccactgagt tggggaaaga  
4800

ggataatcag tgagcactgt tctgctcaga gctcctgac taccccaccc cctaggatcc  
4860

55 aggactgggt caaagctgca tgaaccagg ccctggcagc aacctgggaa tggctggagg  
4920

EP 1 355 149 A2

tgggagagaa cctgacttct ctttccctct cctcctcca acattactgg aactctatcc  
 4980  
 5 tgttaggac ttctgagctt gtttccctgc tgggtgggac agaggacaaa ggagaaggga  
 5040  
 gggtctagaa gaggcagccc ttctttgtcc tctggggtaa atgagcttga cctagagtaa  
 5100  
 10 atggagagac caaaagcctc tgatttttaa tttccataaa atgtagaag tatatatata  
 5160  
 catatatata tttctttaa ttttgagtc tttgatatgt ctaaaaatcc attcctctgc  
 5220  
 15 cctgaagcc tgagtgagac acatgaagaa aactgtgttt catttaaaga tgttaattaa  
 5280  
 atgattgaaa ctg  
 5294  
 20  
 <210> 15  
 <211> 988  
 <212> DNA  
 <213> human  
 25 <400> 15  
 gtcgtgagc gggccttcgc gctggctcgc cgtcggctgc cggggggttg gcctgggtgt  
 60  
 cattggctct gggaagcggc agcagaggca gggaccactc ggggtctggt gtcggcacag  
 120  
 30 ccatggcggg cgcgttggtg cggaagcgg cgactatgt ccgaagcaag gatttcggg  
 180  
 actacctcat gactacgac ttctggggcc cagtagccaa ctgggtctct ccattgctg  
 240  
 35 ccatcaatga tatgaaaaag tctccagaga ttatcagtgg gggatgaca ttgacctct  
 300  
 gttgctatc tttgacattc atgagatttg cctacaaggt acagcctcgg aactggcttc  
 360  
 40 tgtttgcatg ccacgcaaca aatgaagtag ccagctcat ccaggagggg cggttatca  
 420  
 aacacgagat gactaaaacg gcatctgcat aacaatggga aaaggaagaa caaggctctg  
 480  
 45 aagggacagc attgccagct gctgctgagt cacagatttc attataaata gcctccctaa  
 540  
 ggaaaataca ctgaatgcta tttttactaa ccattctatt tttatagaaa tagctgagag  
 600  
 50 tttctaacc aactctctgc tgccttaca gtattaaata ttttactct ttcataaag  
 660  
 agtagctcaa aatatgcaat taatttaata atttctgatg atgttttatc tgcagtaata  
 720  
 55 tgtatatcat ctattagaat ttacttaatg aaaaactgaa gagaacaaaa tttgtaacca



EP 1 355 149 A2

780

ctagcaactta agtactcctg attctttaaca ttgtcttttaa tgaccacaag acaaccaaca  
840

5 gctggccacg tacttaaaat ttgtcccca ctgtttaaaa atgttacctg tgtatttcca  
900

tgcaagtgtat atattgagat gctgtaactt aatggcaata aatgatttaa atatttgtaa  
960

10 aaaaaaaaaa aaaaaaaaaa aaaaaaaaa  
988

<210> 16  
15 <211> 4908  
<212> DNA  
<213> human

<400> 16  
20 ggataacctc gcagggtggg ccggagggcg ggcgccgcg ctgcctgtgc lgcggcgatg  
60

gcccagtgtg tacaatcagt gcaggagcta atcccgact ccttcgtccc ctgtgtcgt  
120

gcgctgtgca gcgacgaagc cgagcggtc actcgtctca atcacctcag cttcgcggag  
180

25 ctgcttaagc ccttctccc cctcacttcc gaggttcaca tgagagatcc taataatcaa  
240

cttcacgtaa ttaaaaattt gaagatagca gtaagcaaca ttgtcaccca gccacctcag  
300

30 cctggagcca tccggaagct tttgaatgat gttgtttctg gcagtcagcc tgcagaagga  
360

ttagtagcta atgtgattac agcaggagat tatgacctta acatcagtgc cactactcca  
420

35 tggtttgagt cttacagaga aacctttctt cagtcgatgc cagcatcgga tcatgaattt  
480

ctgaaccact atttagcatg tatgttggtg gcgtcatcta gtgaagctga acctgtggaa  
540

40 cagttttcaa agttgtcaca agaacagcat cgaattcagc acaacagtga ttattcctac  
600

cccaagtggg ttataccaaa tacacttaa tactatgtac ttttacctga tgtaagtga  
660

45 ggagatgaac agagagctga atcaatttat gaagaaatga aacagaaata tggaactcag  
720

ggttgctatt tacttaaaat taattctcga acatctaata gagcatcaga tgaacagata  
780

50 ccagatcctt ggagtcagta tctccagaaa aatagtattc aaaaccagga atcatatgaa  
840

gatggccctt gtactataac ttcaaataag aattctgata ataacttget ttcattggat  
900

55

ggattagata acgaagtcaa agatggctta ccaaataact ttagagctca ccacttcag  
 960  
 5 ttggagcaat ccagtgaccc ttctaacagt attgatggcc cagatcatct aagatctgct  
 1020  
 tcatcgttac atgaaacaaa gaaaggaaat actggaataa tcatggtgc atgtttaaca  
 1080  
 10 cttactgac atgatagaat tcgacagttt atacaagagt tcacatttcg gggcctttg  
 1140  
 ccacatatag agaaaacaat taggcaatta aacgatcagc taatatcaag aaaaggtttg  
 1200  
 15 agtcgatctc tattttctgc aactaaaaa tggtttagtg gcagtaaagt tccagaaaaa  
 1260  
 agcattaatg acctgaaaaa tacatctggc ttgctgtatc ctccggaagc accagaactt  
 1320  
 20 caaatcagga aaatggctga cttatgtttt ttggtgcagc attatgattt ggcttacagt  
 1380  
 tgctatcata ctgcaaagaa agattttctt aatgatcaag caatgcttta tgcagctggt  
 1440  
 25 gccttgaaa tggcagcagt gtctgctttt cttcaaccag gagcacctag gccatatcct  
 1500  
 gctcattaca tggatacagc aattcagaca tacagagata tctgcaagaa tatggtgttg  
 1560  
 30 gctgaaagat gtgtgttgct tagtgctgaa cttttaaaaa gccaaagcaa atattcagag  
 1620  
 gctgcagctc tctaatacag gttgaccagt gaggattctg atcttcgaag tgcacttctt  
 1680  
 35 ttggaacagg cagcacattg ctttataaac atgaaaagtc ccatggttag aaaatatgca  
 1740  
 tttcatatga tattggcagg ccatcgattt agtaaagcag ggcagaaaaa gcatgcttta  
 1800  
 40 cgctgttatt gtcaagccat gcaagtttac aaaggaaaag gctggtctct tgcaaggat  
 1860  
 cacattaatt tcactattgg gcgccagtcc tatactctta gacagctgga taatgctgtg  
 1920  
 45 tctgctttta ggcataattt aattaatgaa agtaaacaat ctgctgctca acagggggct  
 1980  
 ttctcagag aatatcttta tgtttacaag aatgtaagtc agctgtcacc agatggctct  
 2040  
 50 ttgccacagc ttcctttacc gtatattaac agttcagcaa cacggggttt tttggccat  
 2100  
 gacagacgac cagcggatgg tgaanaacaa gcagctactc atgtaagtct tgatcaagaa  
 2160  
 55 tatgalcttg aatcctctca gcagtggcga gaacttgagg aacaagttgt ttctgtggtt  
 2220

EP 1 355 149 A2

aacaaaggag taattccatc caattttcat cccacacaat actgtttgaa cagttaactca  
2280

5 gataatlcaa gattttccact tgcagttgta gaagaaccaa ttacagtgga agtggttttt  
2340

agaaaccctt tgaaagtctt acttttgttg actgatttgt cattgctttg gaagtttcat  
2400

10 cctaagatt tcagtggaaa ggataatgaa gaagttaaac aactagttac aagtgaacct  
2460

gaaatgattg gagctgaagt tatttcagag ttcttaatta atggcgaaga atcaaaagtg  
2520

15 gcaagactaa agctcttttc ccatacacata ggggagctgc atattctggg agttgtttat  
2580

aatcttggca ctattcaggg ctctatgaca gtagatggca ttggtgctct tcccggatgt  
2640

20 cacacaggaa aatattcctt gagtatgtca gtccgaggga agcaggattt agaaattcaa  
2700

ggtcctcgac ttaacaacac aaaagaagag aaaacatctg ttaaatatgg cctcgatcga  
2760

25 cgtttagatc ccataatcac agaagaaatg ccaactgttg aggtgttctt tatacatttt  
2820

cctacagggc ttctctgttg agaaatccga aaagcatatg tagaatttgt caatgtcagc  
2880

30 aaatgtccac ttactggatt gaaggttgtt tctaaacgtc cagagttctt tactttcggg  
2940

ggtaatactg ctgttctaac accactaagt cctcagctt ctgagaattg tagtgettac  
3000

35 aagactgttg tgacagatgc tacctctgtg tgtacagcac tcatatcatc agcttcttct  
3060

gtagactttg gcattggcac aggaagtcaa ccagaggtga ttctgttcc ccttctcgac  
3120

40 actgttcttc taccgggagc ctcaagtgcag ctgccaatgt ggttacgtgg gctgatgaa  
3180

gaagggtgcc atgaaattaa ctttttgttt tactatgaaa gtgtcaaaaa gcagccaaaa  
3240

45 atacggcaca gaatattaag acacactgca attatttgta ccagtcgggc tttaaatgta  
3300

cgggccactg tctgcagaag taattctctt gaaaatgaag aaggcagagg aggcaatatg  
3360

50 ctagtctttg tggatgtgga aaataccaat actagtgaag caggcgtaa ggaattccac  
3420

ataglgcaag tatcaagtag tagcaaacac tggaagttac agaaatctgt aaatctttct  
3480

55 gaaaacaaag atgccaaact tgccagtagg gagaaggga agttttgctt taaggcaata  
3540

agatgtgaga aagaagaagc ggccacacag tcctctgaaa aatatacctt tgcagatatac  
3600

5 atcttttgaa atgaacagat aataagttca gcaagcccat gtgcagactt cttttatcga  
3660

agtttatctt ctgaattgaa aaaaccacaa gctcacttgc ctgtgcatac agaaaaacag  
3720

10 tcaacagagg atgctgtgag attgattcaa aaatgcagtg aggtagattt gaatattgtc  
3780

atattatgga aggcatacgt tgtggaagac agtaaacagc ttatttttga aggtcaacat  
3840

15 catgttaltc ttcgcactat aggaaaagaa gccttttcat atcctcagaa acaggagcca  
3900

ccagaaatgg aactattgaa atttttcagg ccagaaaaca tlacagtttc ctcaaggcca  
3960

20 tcagtagagc agctttctag tctcattaaa acgagtcttc actaccaga atcatttaat  
4020

catccatttc atcaaaaaag cctttgttta gtaccagtca ctcttttact ttccaattgt  
4080

25 tctaaggctg atgtagatgt catagttgat cttcggcata aaacaacaag tccagaagca  
4140

ctggaaatcc atggatcatt cacatggctt ggacaaacac agtataaact tcaacttaaa  
4200

30 agccaggaga ttcacagtct gcagctgaaa gcatgctttg ttcatacagg tgtttataac  
4260

cttggaactc ctagggtatt tgccaagtta tcggaccaag ttacagtgtt tgaacaagt  
4320

35 cagcagaatt ccatgcctgc cctgatcatc atcagtaatg tgtgacaact tggaaatttg  
4380

tactgaaatc cacaataatc agtttttgct ggatgggttt tacagcagta tttgatatac  
4440

40 ctaacttggt atggagggtg altgatatct gatccctgca aaatactttg acttgtcatt  
4500

ttgttgatga tgcaaagcac gttggactga gaatacttaa cattcttttt ctgtatttct  
4560

45 ttaaaccctg agaataattt acatgctcat aatacaggat atcagcatat ttgtgcacct  
4620

tattaagccc catcttaaga aaacacaaag tctaagtctg ctgttacaac ttgtcaatgg  
4680

50 tatacgaata ttaggagatg attctgagaa aggaaaggcc ttgttggcag tactcctgtt  
4740

aagccattag tctctaaatt ccagctttac tgtgaagttc tatagagtgt taaatacaaa  
4800

55 ttttctgtc ttgcttcaca cagttcctta aaatcagttt tgaactttgg tcatagagtc  
4860

# EP 1 355 149 A2

tlcalatttc agtatttggg ggtccctatg acttatacat aactttgt  
4908

5	<210> 17
	<211> 435
	<212> DNA
	<213> human
10	<220>
	<221> misc_feature
	<222> (30)..(30)
	<223> any kind of base
15	<220>
	<221> misc_feature
	<222> (49)..(49)
	<223> any kind of base
20	<220>
	<221> misc_feature
	<222> (75)..(75)
	<223> any kind of base
25	<220>
	<221> misc_feature
	<222> (76)..(76)
	<223> any kind of base
30	<220>
	<221> misc_feature
	<222> (78)..(78)
	<223> any kind of base
35	<220>
	<221> misc_feature
	<222> (79)..(79)
	<223> any kind of base
40	<220>
	<221> misc_feature
	<222> (109)..(109)
	<223> any kind of base
45	<220>
	<221> misc_feature
	<222> (136)..(136)
	<223> any kind of base
50	<220>
	<221> misc_feature
	<222> (137)..(137)
	<223> any kind of base
55	<220>
	<221> misc_feature
	<222> (149)..(149)

<223> any kind of base

5 <220>  
<221> misc\_feature  
<222> (227)..(227)  
<223> any kind of base

10 <220>  
<221> misc\_feature  
<222> (236)..(236)  
<223> any kind of base

15 <220>  
<221> misc\_feature  
<222> (246)..(246)  
<223> any kind of base

20 <220>  
<221> misc\_feature  
<222> (342)..(342)  
<223> any kind of base

25 <220>  
<221> misc\_feature  
<222> (363)..(363)  
<223> any kind of base

30 <220>  
<221> misc\_feature  
<222> (389)..(389)  
<223> any kind of base

35 <220>  
<221> misc\_feature  
<222> (426)..(426)  
<223> any kind of base

40 <400> 17  
ggtagaaatg attgtgatgt acaaattttn tattttgatc atacttaana agacagagca  
60  
gactcacatt cattnncnna atagtatcac tgtacacata gcgaatttnt ggcgctttta  
120  
gattgctctg aaaatnctg aagagttgnc catagcagcc tggtaacct tttcctttcc  
180  
cccaaagctc tcctgccctt tgcagaaaga ctgttggtga caactgntgc taactnaata  
240  
gcatgnggtt gaacttcgcc aaaatccttc cacctcclcc catagggcaa caggggtgac  
300  
ttgggcttaa agggcattga gtaagcaagt aggttatcag anaacagagg gaagattcca  
360  
ttntagataa tttccaaata ttacaattng tggaactcag agttcaactg ctcaattcct  
420

55

tcttcngctg accct  
435

5

<210> .18  
<211> 2224  
<212> DNA  
<213> human

10

<400> 18  
ctttagatct gtgcagcctt tgcgtgccaa acttgtgaaa ttccttttac cttttttgga  
60

gtacttgcta taaagccacc tgtcaacaaa cccccattat gtacagaata ggacctatcc  
120

15

agtagccagg ccagtaggca gttggggaag gtgggaagga tccagcgagg cccctgagcc  
180

tgcacctgga caggtgtacg tctgcacca tcacctcag caccaggcca ccttgagtc  
240

20

cacttactgt actgtgttgt ggaaggatat gctaagtgat gaaagttgag agcagttcca  
300

ctggtcgtgt aaactttttt ttttttttgg aaattgaagc ttagagtgct tgcccgaat  
360

25

ctctaggaag ttggtggcaa gggacagcac tcacactctt ctggtcatga tctctgatct  
420

ccacctcaaa tgacaataaa aaactgggtcc aacgaagaca ctgctcagca cttcagccat  
480

30

caggactaat ccatcgatga ctggaaaaga ggctagcttt gaggaaaaca gcctgggctc  
540

ttgggagcag agtcacagtgg gtgtgaggct gacttgccga cggtcggcag gtaatggctc  
600

35

tcagccggcg aggcgggtccc acagctctcc tcccagggca gcctgaggag gaggaggccg  
660

ggtgcctggt ttggtggcagc ttcagcctag ggatacctga agctgttgag caacaccttt  
720

40

atgaaatggt gccagagcag caacacttcc ctgtgggcac agccccggga aatccggtac  
780

caagtgaaca aggtggcagg acccacccaa gcctgatacg catctgggcc cgccgggctc  
840

45

agcaggggag gctgctacgg ctgcccactt cccagcaccg tctgtcaggc ttgaacctt  
900

ctgtgctggt cccctcctgg ctaataggga gacccttcgc aggcacccac tgtttcaact  
960

50

tgacctccc acccctgct actctcctcc acacacctt ccgttccgct agcctacctt  
1020

gtcagccttt caataaaagt tatgcacaaa tgtgaacacc tgagatggag ctgaacattt  
1080

55

cttcactttg tcttttttct gaagtcaaac tcttatcaaa tgcctaaaa ttattaccac

1140  
 ccaagagaaa caggaaaaag gttacatgtt ttgttttact gagagtaaga tcacctgcat  
 1200  
 5 ctggaagacg ggctggtaaa ttggtttggc tacagaacag aaagaaaaa aaaacaaacc  
 1260  
 tcgtaaggga agtatcgac tcagacacca ccacttcta gagccaaatg agcaatccca  
 1320  
 10 aactgcaagt gccgtaagtg ggctgtgac gtcacaccgc ccggcccgag gtatcgcatg  
 1380  
 tgcgggggag gccacacta cagctgtcct ctggtctaga aggcaccacc tcgctttcat  
 1440  
 15 gtcccggtgtg ttttgaaaa gcagtatggt gtgtcatgtc tagcggcgaa cacttccctc  
 1500  
 cctctgtcct tgaggttcta atataaaaac tgtgtttctg tacgtgtggg tgggaattct  
 1560  
 20 ctgacggtgc tcgttcatac cacaagctta cgctgagttc lgaactgtcg ttcacagctg  
 1620  
 cgtgtctgca tgggtgtgca tctgtgtac ctttggggaa aatttgtatg taaatgtaca  
 1680  
 25 gaaataaaaa cgttgcccca ttaacagatt tcctctggaa tgtcttcct acctcacctg  
 1740  
 atggtatcca ccgaaggga tttcactacc attaatggtg agtaataaaa tcctccgtgt  
 1800  
 30 tcattcagac ctcaactcgt cactactttg aacgcctctg taagctgtgt cttcacccgc  
 1860  
 cccgaggtgg gtggaggga gcctctcact ctgcttcgag tcctggtctt aaaggtagtc  
 1920  
 35 agaggcagag gctggattaa acacacactg ttaccaagt gccactctca gaccacctga  
 1980  
 gagacgggg gccatcagta aaattaagag gaattttttt cccttggtcg tgtatgttct  
 2040  
 40 gctgatccgt ggctgaagg ttcctagaga cgtcaagaaa tgaatatctt aactgtgat  
 2100  
 tctgtgagga aagactgga acccaaaact ctcttctcta atgtattttt taacgaaaat  
 2160  
 45 gacaatatct ctttaataaa gtatttatac caaaaaaaaa aaaaaaaaaa aaaaaaaaaa  
 2220  
 aaaa  
 2224  
 50  
 <210> 19  
 <211> 2244  
 <212> DNA  
 <213> human  
 55  
 <400> 19



EP 1 355 149 A2

gttggtttaa agcaaggcat gcttggtgat gactctgtaa cagactaatt ggaattgttg  
60

5 aagctgctcc ctggttccac tctggagagt aatctgggac atcttagtgt tttgttttgt  
120

ttttttccct cctctttttt tgggggggag tgtgtgtggg gtttqttttt tagtcttgtt  
180

10 tttttaattc attaaccagt ggttagcctt aaggggagga ygacggattg attccacatt  
240

ccacttccta gatctagtlL agaaaacatg ttcccatctt ggtgctctta ggaaggagta  
300

15 tagtaaatgc ctcatTTaat aacatactcc tttttgaaag ttgccttttc tctccacctt  
360

tgagtagatc cagtatttga tgaactcat gaaagtgggt ggagcccatc tccccctcc  
420

20 tcttttctag gacgcactat atgtgactgt gactttaagg acatttgttt gccatttgct  
480

gattttttlg ggaagttaat ttctaacttc tttcactgat aaatgaagaa aagtattgca  
540

25 cctttgaaat gcaccaaag aattgagttt gtaattaaaa aaattttttt tccctttcag  
600

tcattgtctt atatgcttag catagatttg cagctcagta gtatatgtgt tcctagaatg  
660

30 cagctgaaga cctgttatgt agaggaaata cgaggggtgg tgctagaaga cagacatctg  
720

tggaatgatt cacatcctct caagttagga ggatggaggc ctgcttcatt aagaagctgg  
780

35 gggtaggggtg ggggtgggga gaacacttaa caacatgggg accagtcagg ggaatcccc  
840

tatttctgtt ttgcatatga ggaaccctag agcagccagg tgaggctctc tagtttaata  
900

40 aaaatcatgg aaagacLctt aatgcagact cttcttaagt gttaataggg attttttcag  
960

cttatttttg ttgcagtttc caatttttaa aaatgttgag gtaatctttc ccaccttccc  
1020

45 aaacctaatt cttgtagatg cattagtgtt gaaccaatgc ttctcatgtc tcaatcttgt  
1080

atatcatctt ttcagatgta ttaacaaaca aaaccttaaa aagagtagat gaattgccaa  
1140

50 acacaattcc taccaataat aaatcgatca actctatcta ttcaggaaag caggaagcat  
1200

ttggaccaca gtgcatgaaa acttcaacat tctgttatta gataatgaat caaccaaatg  
1260

55 aacaatccag agaaaagaaa attgcaataa taaaaggtaa attaacagaa agataatata  
1320

agcaagatag taatagttga ccattctgaa aagcttataa catcactcat catccagcat  
 1380  
 5 cctttctgaa aacaaaggat ttttaaatca ctttatgcac atatacaaca taggaggttg  
 1440  
 gcaaaataat gcactatttc ttaacagcca tgtctcttgt agaacttcaa gttaatctac  
 1500  
 10 aaatgaccat tgtgtcttaa tttagattat gaalaccaca ttagtcaggt atttgacta  
 1560  
 acccttaata gtatatacag tttctatgga aaattcagtg gtccaaaaat ttccgtagaa  
 1620  
 15 tttgagagga cgttggtggg ctgaagatag ctccttgagg gtcactgatg taggctgcaa  
 1680  
 tgggggttca caaggccctg acaccgtatt tatagtctaa cctttttatg aaaatctgac  
 1740  
 20 tacagctatt taaggagtag tcttaatagc tgaaaatgaa gatagagaaa gacaccaaga  
 1800  
 atatgacaca gtttacattc tagtgaggga cacaacaaaa tcaaatttaa aaaagagtgt  
 1860  
 25 aatagatgct gataaatact gtagataaag cacataagaa aatagaaata aaggctgtca  
 1920  
 atggagaagt catgattttt attttattta tttatttatt tatttgagac agagtcaggc  
 1980  
 30 tctgtgcagg ctggagtgca atggtgtgat ctcgctcact acaacctctg ctcctggctc  
 2040  
 aagctatcct cccacctcag ctctcaagta gctgggatca cagggtgcgtg ctaccatgcc  
 2100  
 35 cggctaattt tttgtagaga tgagggtttg ccatgttgcc caggctgggc tcgaactcct  
 2160  
 ggactcaact gacccccact cggcctctca aagtgctgag attataggcg tgcagccggc  
 2220  
 40 agctggccat tgtttatggt ctgc  
 2244  
 45 <210> 20  
 <211> 351  
 <212> DNA  
 <213> human  
 <220>  
 <221> misc\_feature  
 <222> (62)..(62)  
 <223> any kind of base  
 50 <220>  
 <221> misc\_feature  
 <222> (121)..(121)  
 <223> any kind of base  
 55

<220>  
 <221> misc\_feature  
 <222> (207)..(207)  
 <223> any kind of base  
 5

<220>  
 <221> misc\_feature  
 <222> (220)..(220)  
 <223> any kind of base  
 10

<220>  
 <221> misc\_feature  
 <222> (276)..(276)  
 <223> any kind of base  
 15

<220>  
 <221> misc\_feature  
 <222> (300)..(300)  
 <223> any kind of base  
 20

<220>  
 <221> misc\_feature  
 <222> (315)..(315)  
 <223> any kind of base  
 25

<220>  
 <221> misc\_feature  
 <222> (336)..(336)  
 <223> any kind of base  
 30

<400> 20  
 tctacttcca catcgcgag accgagaagc gctgtttcat cgaggaaatc cccgacgaga  
 60  
 35  
 cnatggtcat cggcaactat cgtaccaga tgtgggataa gcagaaggag gtcttctctgc  
 120  
 nctcgacccc tggcctgggc atgcacgtgg aagtgaagga ccccgacggc aaggtggtgc  
 180  
 40  
 tgtcttgga gtacggctcg gagggcnctt tcacgttcan ctcccacacg cccggtgacc  
 240  
 atcaaatctg tctgcactcc aattcttacc aggatngctc tctttcgctg gtgggcaaan  
 300  
 45  
 tgcgtgttgc atctngacat ccaggttttg gggagnatgc caacaaatta c  
 351

<210> 21  
 <211> 2631  
 <212> DNA  
 <213> human  
 50

<400> 21  
 accttccaac ccagccctcg gctgagccgc gccgcaccat gcccgccgtg gacaagctcc  
 60  
 55  
 tgctagagga ggcgttgca gacagccccc agactcgctc tttactgagc gtgtttgaag

120  
aagatgctgg caccctcaca gactatacca accagctgct ccaggcaatg cagcgcgtct  
180  
5 atggagccca gaatgagatg tgcctggcca cacaacagct ttctaagcaa ctgctggcat  
240  
atgaaaaaca gaactttgct cttggcaaag gtgatgaaga agtaatttca acactccact  
300  
10 atttttccaa agtgggtgat gagcttaatc ttctccatac agagctggct aaacagttgg  
360  
cagacacaat ggttctacct atcatacaat tccgagaaaa ggatctcaca gaagtaagca  
420  
15 ctttaaagga tctatttga ctcgctagca atgagcatga cctctcaatg gcaaaatata  
480  
gcaggctgcc taagaaaaag gagaatgaga aggtgaagac cgaagtcgga aaagaggtgg  
540  
20 ccgcgcccg gcggaagcag cacctctcct cccttcagta ctactgtgcc ctcaacgcgc  
600  
tgcagtacag aaagcaaatg gccatgatgg agcccatgat aggccttgcc catggacaga  
660  
25 ttaacttttt taagaaggga gcagagatgt tttccaaacg tatggacagc tttttatcct  
720  
ccgttgca ga catggttcaa agcattcagg tagaactgga accgagggcg aaaagatgcg  
780  
30 ggtgtcccag caagaattac tttctgttga tgaatctgtt tacactccag actctgatgt  
840  
ggccgcacca cagatcaaca ggaacctcat ccagaaggct ggttacctta atcttagaaa  
900  
35 caaaacaggg ctggtcaccg ccacctggga gaggccttat ttcttcaccc aaggcgggaa  
960  
tctcatgtgt cagcccaggg gagccgtggc tggagglttg atccaggacc tggacaactg  
1020  
40 ctcatgtatg gccgtggatt gcgaagaccg gcgctactgc tttcagatca ccacgcccac  
1080  
tggaaaaatcg ggaataatcc tccaggctga gacagaaaag gaaaatgaag agtggatatg  
1140  
45 tgcaataaac aacatctcca gacagatcta cctgaccgac aacctgagg cagtcgcat  
1200  
caagttgaat cagaccgctc tgcaagcagt gactcctatt acaagttttg gaaaaaaca  
1260  
50 agaaagctca tgccccagcc agaacctgaa aaattcagag atggaaaatg aaaatgacaa  
1320  
gattgttccc aaagcaacag ccagtctacc tgaagcagag gagctgatcg cgctggagc  
1380  
55 gccgattcaa ttcgatattg tgcttctgc tacagaattc cttgatcaga acagagggag

EP 1 355 149 A2

1440  
caggcgtacc aacccttttg gtgaaactga ggatgaatca tttccagaag cagaagattc  
1500  
5 tcttttgcag cagatgttta tagttcgggtt ttgggatca atggcagtta aaacagacag  
1560  
cactactgaa gtgatttatg aagcagatgag acaagtattg gctgctcggg ctattcataa  
1620  
10 catcttccgc atgacagaat cccatctgat ggtcaccagc caatctttga ggttgataga  
1680  
tccacagact caagtatcaa gggccaattt tgaacttacc agtgtcacac aatttgctgc  
1740  
15 tcactcaagaa aacaagagac tggttgggtt tgtcatccgt gttcctgaat ccactggaga  
1800  
agaatctctg agtacatata tttttgaaag caactcagaa ggcgaaaaga tatgttatgc  
1860  
20 tattaatttg ggaaaagaaa ttattgaggt tcagaaggat ccagaagcac tggctcaatt  
1920  
aatgctgtcc ataccactaa ccaatgatgg aaaatatgta ctgttaaacc atcaaccaga  
1980  
25 tgacgatgat ggaaatccaa atgaacatag aggcgcagaa tccgaagcat aactcacttg  
2040  
cgcctgtggg ggaagagcga acaggaagga gagctacctc ctaaggggtt taactgtctc  
2100  
30 gacatacagg cacactgacc tgatttccga aggcctgacaa tcgtttgtgg aatgtaatct  
2160  
tgatgccttg atactgagac ttgggagggg aactaagaaa tggttgacag cgttcccacc  
2220  
35 catctacaat gttatttttag gtgctttgtg gtaagtcttt tttcttagat tgcgctaaaa  
2280  
tttcttagat tgttcagcgc tcagaacaaa agtttgaaaa atgcattggt catatgaatg  
2340  
40 tcactctctt tcagtttcca gtatcctttt taataaatgg caaaagccta gatttacaat  
2400  
ttgatgaaca ctaaataatt cttattaata taatctattt ttgtatttta cttaatgagc  
2460  
45 tttaagtgcc tgtcgttctg aaaattgtgt atttataatt cagcttatcl cataattgga  
2520  
cctaatagca tttctttgtg cagttaggtg atgagcactg ctttgaggcc caagcactag  
2580  
50 tagagatgcg cgatacaggt ctagtctcgg taactgttcc agacatcaag c  
2631  
<210> 22  
55 <211> 2851  
<212> DNA

EP 1 355 149 A2

```

<213> human
<400> 22
5 agcatctcag gccatcatcc tgaaacttgg cagccttcgt ggagtataag gacagcatta
60
   ttagccatca ttgggtttac tgccaacaaa aggagaggga gccatagggt ctctagatta
120
   cactcctgag gaaagaagag cacttgccaa aaaatcacia gatttctggt gtgaaggatg
10 180
   tggctctgcc atgaaggatg tcctgttqcc tttaaaatct ggaagcgatt caagccaagc
240
   tgaccaagaa gccaaagaac tggctaggca aataagcttt aaggcagaag tcaattcatc
15 300
   tggaagact atctctgagt cagacttaaa ccactctttt tcaactaactg atttacaaga
360
   tgatatacct acaacattcc aggggtgtac ggccagtaca tcgtacggac tccagaattc
20 420
   ctccagcaga tcctttcacc aacctaccca acctgtagct aagaatacct ccatgagccc
480
   tcgacagcgc cgggccagc agcagagtca gagaagggtg tctacttcac cagatgtaat
25 540
   ccagggccac cagccaagag acaaccacac tgatcatggt gggtcagctg tactgattgt
600
   catcctgact ttggcattgg cagctcttat attccgacga atatctctgg caaacgaata
30 660
   catatttgac ttgagttat aatatggttt tgtgacttat gagctgtgac tcaactgctt
720
   cattaacat tctgcattgg gtataatcta agaattgttt acaaaaagat tattttgtat
35 780
   ttacccttca ttctttttt tgatccttgt aagtttagta taaatatatc tagacattca
840
   gactgtgtct agcagttacg tcctgcttaa agggactaga agtcaaagtt ccttglctca
40 900
   ctatttgatc tgctttgcag ggaaataact tgttttttct catgtttcat cttcttttta
960
   tgtaaatttg taatactttc ctataattgcc ctttgaaatt ttggataaa agatgatgtt
45 1020
   ttaagttcca atgagtatta ctagttactc aataccactt attgagtact ctgtttctac
1080
   gtatgtagaa tgtataggga tagaagagtt gaaaaggga agcaaaactt ctttaagtggc
50 1140
   ttccttaaaa tgtcattcat aggagatgta ctggaattgc tcattctgtg actttatttg
1200
   tgtcctaaac attcttcagt gaaaataatt ttatttcagt caaacattta tgaggaaatg
55 1260

```

EP 1 355 149 A2

agatcacatc tttgtcactg gatgtactt gaagagggag tactttgtaa ccactttgat  
1320

5 atgctgttat caccaccccc tgccctctgc tgccataatc acacaaattt aaaaagaaag  
1380

aaaacagtct tccatagatt ttttaaggaag aaagggccca agtcaggaga tcgcttggtt  
1440

10 ttcttcaga agttaaatgg ggggatctga agatttgaat gttcggctcg ctttgaaatg  
1500

tatgtctttt gggaaatgat tatatgccta gctttataat caggtataaa attttaatta  
1560

15 ttcccaggaa tatgcataat attgaatatt tcatgtccta ttttaataga aaacctcagg  
1620

gcccagtaa ccagtgatag aagttagaaa aacctcttta cttagaattg tccacctagt  
1680

20 cagagcccaa gaaagaattt tcagtygaaa aatcaatata taacttagtg ctagctagcg  
1740

ccacagactc tagtagataa tattatcatc ataatggctg gtgaaaccat ataatcacag  
1800

25 aaaaacattg ccttcagcat gttcagttcg cagcactgag ggcactcttg aggggtgtgt  
1860

taatgaagat ttaattttta aatacaggty gttccaagct ttcaaatagg ttatgtctca  
1920

30 aaagtgttat ttgtaagtta atttttttac aagtcaaaca atgttggaag tgggtatttag  
1980

gttctagatc ggtccacgaa agttagccca tatgtatatc ttgaatagta tagggggagg  
2040

35 tattcataaa gtccttatgt ggttttaact aagtgaatt atggacaaga gaaataattg  
2100

taaaatcgtc ttaaaggcaa atttaatttt taccctgtt tatgggacat tcgttctatt  
2160

40 aactgtcaga cacaaattct gttttcatct gagagccagg tttcctttat ttctacatct  
2220

aaaataagaa catattgtac actattatat aatacagaat tgtcttacac tttataaat  
2280

45 tcgcatttta aagtggttta caggattatt ttttatatct gtagctgaat ttgttaaagt  
2340

ctaaaaagct caaggacttt atgaagatct cattatatga ggaaatcat aggttaccat  
2400

50 ttataactc tattgccata agaaaataca ctctaaaatc ttgatttgaa acatattaga  
2460

aaaccttgatt cagtgtcag tgggtctcta gtaagaagtc accgacggta gcgtcatatg  
2520

55 agaagaaaga aatccccacc acctcaacct ctgctgagat tgtgtgctag gaacagcctt  
2580

EP 1 355 149 A2

ccctccgttt cccctcagtc aaacttgagc cagcctctgg atcgatgtga tcttattgca  
2640

5 tgtttccatg ggggtgtacct atactttaag ccaatcctgc tgcattcact gctaagttaa  
2700

ataaaaagcc aagaagaaaa aaaaaatttt gcactgtgca gatcctttgc tatctgactt  
2760

10 gcactctcttc cccacactgt cagctagcca cctgcttgtt tgtgttggga ttttttttag  
2820

cacctgaagc accatctgaa aggggcacca t  
2851

15 <210> 23  
<211> 3473  
<212> DNA  
<213> human

20 <400> 23  
aagagcagcg gcgaggcggc ggtggtggct gaggcctgg tggcagaggc gaaggcgaca  
60

gctctagggg ttggcaccgg ccccgagagg aggatgcggg tccggatagg gctgacgctg  
120

25 ctgctgtgtg cgggtgctgt gagcttggc tcggcgtcct cggatgaaga aggcagccag  
180

gatgaatcct tagattccaa gactactttg acalcagatg agtcagtaaa ggaccacact  
240

30 actgcaggca gagtagttgc tggtaaata tttcttgatt cagaagaatc tgaattagaa  
300

tcctctattc aagaagagga agacagcctc aagagccaag agggggagag tgtcacagaa  
360

35 gatatcagct ttctagagtc tccaaatcca gaaaacaagg actatgaaga gccaaagaaa  
420

gtacggaaac cagctttgac cgccattgaa ggcacagcac atggggagcc ctgccacttc  
480

40 ccttttcttt tcctagataa ggagtatgat gaatgtacat cagatgggag ggaagatggc  
540

agactgtggg gtgctacaac ctatgactac aaagcagatg aaaagtggg cttttgtgaa  
600

45 actgaagaag aggctgctaa gagacggcag atgcaggaag cagaaatggt gtatcaaact  
660

ggaatgaaaa tccttaatgg aagcaataag aaaagccaaa aaagagaagc atatcggtat  
720

50 ctccaaaagg cagcaagcat gaaccatacc aaagccctgg agagagtgtc atatgctctt  
780

ttatttgggtg attacttgcc acagaatatc caggcagcga gagagatgtt tgagaagctg  
840

55 actgaggaag gctctcccaa gggacagact gctcttggct ttctgtatgc clctggactt  
900



ggtgttaatt caagtcaggc aaaggctctt gtatattata catttgagc tcttggggc  
 960  
 5 aatcataatg cccacatggt ttgggttac agatactggg ctggcatcg cgtcctccag  
 1020  
 agttgtgaat ctgccctgac tcaactatcgt cttgttgcca atcatgttgc tagtgatato  
 1080  
 10 tcgctaacag gaggtcagc agtacagaga atacggctgc ctgatgaagt ggaaaatcca  
 1140  
 ggaatgaaca gtggaatgct agaagaagat ttgattcaat attaccagtt cctagctgaa  
 1200  
 15 aaagtgatg tacaagcaca ggttggctt ggacaactgc acctgcacgg agggcgtgga  
 1260  
 gtagaacaga atcatcagag agcaattgac tacttcaatt tagcagcaaa tgctggcaat  
 1320  
 20 tcacatgcca tggcctttt gggaaagatg tattcggaag gaagtgcacat tgtacctcag  
 1380  
 agtaatgaga cagctctcca ctactttaag aaagctgctg acatgggcaa ccagttgga  
 1440  
 25 cagagtgggc ttggaatggc ctaccttal gggagaggag ttcaagttaa ttatgatcta  
 1500  
 gcccttaagt atttcagaa agctgctgaa caaggctggg tggatgggca gctacagctt  
 1560  
 30 ggttccatgt actataatgg cattggagtc aagagagatt ataacaggc cttgaagtat  
 1620  
 tttaattag cttctcaggg aggccatato ttggctttct ataacctagc tcagatgcat  
 1680  
 35 gccagtggca ccggcgtgat gcgatcatgt cacactgcag tggagltgt taagaatgta  
 1740  
 tgtgaacgag gccgttggtc tgaaaggctt atgactgcct ataacagcta taaagatggc  
 1800  
 40 gattacaatg ctgcagtgat ccagtaccto ctctggctg aacagggcta tgaagtggca  
 1860  
 caaagcaatg cagcctttat tcttgatcag agagaagcaa gcattgtagg tgagaatgaa  
 1920  
 45 acttatccca gagctttgct acattggaac agggccgcct ctcaaggcta tactgtggct  
 1980  
 agaattaagc tcggagacta ccatttctat gggtttgga ccgatgtaga ttatgaaact  
 2040  
 50 gcatttatcc attaccgtct ggcttctgag cagcaacaca gtgcacaagc tatgtttaat  
 2100  
 ctgggatata tgcatgagaa aggactgggc attaaacagg atattcacct tgcgaaacgt  
 2160  
 55 ttttatgaca tggcagctga agccagccca gatgcacaag ttccagctt cctagccctc  
 2220

tgcaaattgg gcgtcgtcta tttcttgag tacatacggg aaacaaacat tcgagatatg  
 2280  
 5      ttcacccaac ttgatatgga ccagcttttg ggacctgagt gggaccttta cctcatgacc  
 2340  
 atcattgcgc tgctgttggg aacagtcata gcttacaggc aaaggcagca ccaagacatg  
 2400  
 10      cctgcaccca ggcctccagg gccacggcca gctccacccc agcaggaggg gccaccagag  
 2460  
 cagcagccac cacagtaata ggcactgggt ccagccttga tcagtgcag cgaaggaagt  
 2520  
 15      tatctgctgg gaacacttgc atttgattta ggaccttga tcagtggta cctcccagaa  
 2580  
 gaggcacggc acaaggaagc attgaattcc taaagctgct tagaatctga tgcctttatt  
 2640  
 20      ttcagggata agtaactctt acctaaactg agctgaatgt ttgtttcagt gccatatgga  
 2700  
 ataacaactt tcagtggctt tttttttct tttctggaaa catatgtgag aactcagag  
 2760  
 25      taatgtctac tgtatccagc tatctttctt ggatcctttt ggtcattatt tcagtgtgca  
 2820  
 taagtcttta atgtcaacca tctttaaggt attgtgcac gacactaaaa actgatcagt  
 2880  
 30      gtaaaaagga aaaccagtt gcaagtttaa acgtgttcga aagtctgaaa atagaacttg  
 2940  
 ccttttaagt taaaaaaaa aaaagctatc ttgaaaatgt tttggaactg cgataactga  
 3000  
 35      gaaactctta ccagtcacac tgcaattaga catattcagc atatttgta ttttaaagg  
 3060  
 gagggttggg aggtttctta ttggtgattg tcacacgga taccatactc ctctcctca  
 3120  
 40      aagaatgaaa ggccttgta aggagttttt tgtgagcttt acttcttttg aatggaatat  
 3180  
 acttatgcaa aaccttgta actgactcct tgcactaacg cgagtttgcc ccactactc  
 3240  
 45      tgtaatttgc ttgtttgttt tgaatataca gagccttgat ccagaagcca gaggatggac  
 3300  
 taagtgggag aaattagaaa acaaacgaa ctctggttgg ggtactacga tcacagacac  
 3360  
 50      agacatactt ttcctaaagt tgaagcattt gttcccagga tttattttac ttgcatctc  
 3420  
 cttttgcaca aagaacacat caccatttcc ttttgacaa agaacacac acc  
 3473  
 55      <210> 24

EP 1 355 149 A2

```

<211> 401
<212> DNA
<213> human

5      <220>
      <221> misc feature
      <222> (252)..(252)
      <223> any kind of base

10     <220>
      <221> misc feature
      <222> (303)..(303)
      <223> any kind of base

15     <220>
      <221> misc feature
      <222> (390)..(390)
      <223> any kind of base

20     <400> 24
      ttagattatt ttcaalltat tattcagaat aaatatatct tttttcttta acttctcaaa
      60

      tagttattga attgtattgg tttaaattaa atgcgtcatg tgtatatatc agtattaatt
      120

25     caagagatac aaaaggaaat tgagtgaaaa ataagctctgc ctccctccca tcaactctcat
      180

      gtctctacct agaggcaatt attgtcaaca gtttttgatg tgtctttcaa aaaatagtcc
      240

30     attaagcctg gngtactaga tctcttttaa aagtttaca cctgttacag aatatatata
      300

      aangttcaat tactagtaac accttattac agatacagat tacaacttag gaaatatatt
      360

35     ttcattggacc attgatgtca ttgggattcn cccctacaat c
      401

40     <210> 25
      <211> 1820
      <212> DNA
      <213> human

      <400> 25
      aatgtcttag aaaaaggctt tctaaaagaa aaagagcaag aggccatttc ttttcaagat
      60

      agatacaaag aacttcagga aaaacataaa caagaattgg aagacatgag gaaagctggg
      120

50     cacgaagccc tcagcattat tgtggatgaa tataaggcac tactgcagtc ttcagttaag
      180

      caacaagtag aagctattga aaaacagtac atttctgcaa ttgagaaaca ggcacacaag
      240

      tgtgaggagt tgctaaatgc tcagcatcag aggctccttg aaatgctaga tacagagaag
      300

55

```

EP 1 355 149 A2

gaactgttaa aagaaaaaat aaaggaagct ttgattcagc aatctcaaga acagaaggaa  
360

5 atattggaaa agtggttgga ggaagaaagg caaagaaata aagaggcatt agtatccgct  
420

gcaaagcttg aaaaagaagc agtgaaggat gcagttttaa aagtcgtaga agaagaaaga  
480

10 aaaaatttag aaaaagcgca tgctgaagaa aggaattat ggaagacaga acatgcaaaa  
540

gatcaagaaa aagtatctca ggaaaltcaa aaagctatac aagaacaaag aaaaataagt  
600

15 caggaaactg ltaaggcagc aataatagaa gagcagaaac gaagtgaana ggctgtggaa  
660

gaggcagtga aaagaacaag agatgaattg atagagtata taaaagaaca gaaaaggctc  
720

20 gatcaagtca tccgccaaag aagcctgtcc agtttggaac tgttcctctc ctgtgcacag  
780

aaacagttaa gtgctttaat agctacggaa ccagttgaca ttgaataaaa agaacatgac  
840

25 aaacccacac tggcattgga taaatcatat tacaccttca aaatacacac tctgaattat  
900

aaagatgtgt ttgttttctt tccaaatcat gtagaattga tttccagttc aaggataaac  
960

30 caaacaata tttagaacta tcaagtgatc taatttattt tcttttggtt tcttctttac  
1020

atttactgtt attttattat tattagtagt agcagcaaca gagtatgata tgacccaaaa  
1080

35 gccattgtaa agtgccacat taccaaaatt aattaagtaa actttatagc ctgtgggagt  
1140

ctattatata ttattttgca aaagtagtaa atatattatt gtttcatgat gactcttgat  
1200

40 gagatgctag aatgtaacca tacatttalc ttattttgag gatagaaata gcatggattt  
1260

caacatcact talttatctg tataattgga aataaaacac cgatatgata gagaatcatt  
1320

45 ccggcattac ctaacctctt ctgcagttgg atctatgtat tttcattggt ctactgaaaa  
1380

cgaacaatac aattaaaagc actaaagatt attatattaa ttcaactttg atctgatata  
1440

50 tcaactaaac taaaggggtg tgtgtggtgt atgcttggtt cctatttctg ctctttaaag  
1500

atactttgaa tcaataaaac cattagtcta caaatcaaat tgtgaactta atctctagaa  
1560

55 agagaatata actcagccat ttataggaat ttaggttcaa gtacaggata tatgaaatct  
1620

EP 1 355 149 A2

tttcccagta tttcagaatg tacttaattc acaggcagga tgcttcaatg caaaatcatg  
1680

5 aatattttta attcaaaact aaaatgtcat taatatgtat gtatgcaaat gttttatctt  
1740

atttttcgaa atgcatctac tttcatgggc tttgtacgtt tctgagattt ctcagtgtaa  
1800

10 taaaaagagc tcccaaactt  
1820

15 <210> 26  
<211> 280  
<212> DNA  
<213> human

<220>  
<221> misc\_feature  
<222> (261)..(261)  
<223> any kind of base

20 <220>  
<221> misc\_feature  
<222> (237)..(237)  
<223> any kind of base

25 <400> 26  
tcaagtcata agataaagtt taatcatttg atcatgttaa aagacacaaa acacagccaa  
60

30 tctaaccaaa tttcaggcat gcatttacat aaatatatta aattaagaaa agaaattgta  
120

cacttaaacg tccttttcac ctagaatca ttaaatccac agatcaacaa taaaaccaat  
180

35 tctctgcatt taccacttca agatacaatt gttctatttt aaagataaca caaactncac  
240

tagtctgggt aggaatttat ntgcattata catatattat  
280

40 <210> 27  
<211> 392  
<212> DNA  
<213> human

45 <400> 27  
ttggtttgaa atggcacccc aggactttgg gcctgcctta cttgatagcc tcgttcagtg  
60

agcaaagact tagtgagcag ctcttgatg ccaagtattt tgctaagctc tggaaaaaag  
120

50 ataaacaaga catggttctt gctttcaagg agtgtgtaat tcttttagcca gatatggaaa  
180

cctggaccct gagtgggaga aaggagacag atgaaaggag tccgtgattt tgtaaccaag  
240

55 agctgcctgc atggttatga gtatcactga ttttagggac gccacagag ctaaagcatt

EP 1 355 149 A2

300  
 tttttaatcc gagaagactt ttgtaactca tattagttaa tcttctagct ctgagatagc  
 360  
 5 aacacagctc ttagaattct gtaagtaagc tt  
 392  
 <210> 28  
 10 <211> 2299  
 <212> DNA  
 <213> human  
 <400> 28  
 15 cgaacccccca cagctggagg gcgaggccag ctgtaccggg cccagtgcc ctttcgaggc  
 60  
 cacaagcggc cgtctcctg gtccqgtgct ccggcgctg atctagggtc atggagccgg  
 120  
 ggctgtggct ccttttcggg clcacagtga cctccggcgc aggattcgtg ccttgctccc  
 180  
 agtctgggga tgcctggcagg cgcggcgtgt cccaggcccc cactgcagcc agatctgagg  
 240  
 gggactgtga agagactgtg gctggccctg gcgaggagac tgtggctggc cctggcgagg  
 300  
 ggactgtggc cccgacagca ctgcagggtc caagccctgg aagccctggg caggagcagg  
 360  
 cggccgaggg ggcccclgag caccaccgat ccaggcgctg cacgtgcttc acctacaagg  
 420  
 30 acaaggagtg tgtctactat tgccacctg acatcatttg gatcaacact cccgaacaga  
 480  
 cggtgcccta tggactgtcc aactacagag gaagcttccg gggcaagagg tctcgggggc  
 540  
 35 cacttcagg gaatctgcag ctctcacalc ggccacactt gcgctgcgct tgtgtgggga  
 600  
 gatatgacaa ggccctgcctg cacttttgca cccaaactct ggacgtcagc agacagggtg  
 660  
 40 aagtcaagga ccaacaaagc aagcaggctt tagacctcca ccatccaaag ctcatgccg  
 720  
 gcagtggact cgccctcgct ccattctacct gcccccgctg cctctttcag gaaggagccc  
 780  
 45 cttaggagga caggcctgca gcatcctggt ctggggaggc ttctgtcatt gctcacacac  
 840  
 agttcagatt tccacctctt tatagacaag aagtgaattt gcctggggca gaacaccac  
 900  
 50 ccaaagagtc ccacttaac aatacccccc cccacggca agaagccca aatccgaatg  
 960  
 accccagttt tcctaatgag taaaatgatc ccagatgtgc cccagagcat gacgcctgca  
 1020  
 55

EP 1 355 149 A2

gctccggttt catgcaggaa attggttttg gagagtlttg gcaagttgga aagccactta  
1080

5 ctggcttttg acatgacttc tcttgagaa taagtggact ccaagctaac tctttgcaaa  
1140

tgtaaacaca tgtccatctt gtaataaatg caaaatgcc gtgcagcaga agcatgcgac  
1200

10 tttcatatcc ttgcctagaa taggctgcat ggtgtatgtc agtgagggcc acgaggcgct  
1260

ggcttttagac acagatcata gctctacagg agtttatgaa ttgaaagctt atgggatttt  
1320

15 ggagagaaaa ttttcagctg tgcttgatac ccaccaaaag aatgtatctc gaaagaatga  
1380

aggaagaaga aaaaaggatc cttgatgttt gtgacaagaa aatgagaaag ttagtatctg  
1440

20 caatacagag cttgttcctg ttcagtgact gacctctgt attctgtata gacaccaggc  
1500

cgatacacag tggagttccc aggccttgtt tgcaggaagc cgactgtaaa gacagcccca  
1560

25 gctcaaggct attaggttga atatttgctt tcatgagtaa atgtggatct ttggggaatg  
1620

gcttcaaaat aagtcaagaa cacaaattct ttgtaaatta tgtaaattcc tgtttatata  
1680

30 aattggcaac aacttatacc gtctgacagt tcaaaatctc tttcagctgc gctcttccca  
1740

ccgagccgag cttactgtga gtgtggagat gttatccac catgtaaagt cgcctgcgca  
1800

35 ggggagggtt gccatctcc ccaacccagt cacagagaga taggaaacgg catttgagtg  
1860

gggtgccagg gccccgtaga gagacattta agatgggtga tgacagagca ttggccttga  
1920

40 ccaaatgtta aatcctctgt gtgtatttca taagttatta caggataaaa agtcatgacc  
1980

tatcatgagg aaatgaaagt ggctgatttg ctggtaggat tttgtacagt ttagagaagc  
2040

45 gattatttat tgtgaaactg ttctccactc caactccttt atgtggatct gttcaaagta  
2100

gtcactgtat atacgtatag agaggtagat aggtaggtag attttaaatt gcattctgaa  
2160

50 taaaaactca tactccttag agcttgaatt acatttttaa aatgcatatg tgctgtttgg  
2220

caccgtggca agatgggtatc agagagaaac ccatcaattg ctcaaatact cagaaagtac  
2280

55 tgtcaaaagc ctaataaaa  
2299

<210> 29  
 <211> 1339  
 <212> DNA  
 <213> human

5

<400> 29  
 ctaaacaaaa tcattcactt ccctgatttt gataagaaaa ttcctgtaaa gctgtttcct  
 60

10

ctgcctctcc tctacgttgg aaaccacata agtggattat caagcacaag taaattaagc  
 120  
 ctaccgatgt tcaccgtgct caggaaattc accattccac ttaccttact tctggaaacc  
 180

15

atcatacttg ggaagcagta ttactcaac atcatcctca gtgtctttgc cattattctc  
 240  
 ggggctttca tagcagctgg gtctgacctt gcttttaact lagaaggcta tatttttgt  
 300

20

ttcctgaatg atatcttcac atcagcaaat ggagtttata ccaaacagaa aatggacca  
 360  
 aaggagctag ggaatacgg agtacttttc tacaatgcct gcttcgatgat tatcccaact  
 420

25

cttattatta gtgtctccac tggagacctc caacaggcta ctgaattcaa ccaatggaag  
 480  
 aatgttgtgt ttactctaca gtttcttctt tctgtttttt tggggtttct gctgatgtac  
 540

30

tccacggttc tgtgcagcta ttacaattca gccctgacga cagcagtggg tggagccatc  
 600  
 aagaatgtat ccgttgccca cattgggata ttaatcgggtg gagactacat tttctctttg  
 660

35

ttaaaacttg tagggtaaa tatttgcatt gcagggggct tgagatattc ctttllaaca  
 720  
 ctgagcagcc agttaaaacc taaacctgtg ggtgaagaaa acatctgttt ggatttgaag  
 780

40

agctaaagag tctgcagcag gattggagac tgacttgtga ctgcgggctg ggggggcatt  
 840  
 cccagtagga atgtgaagcc agagggttcg gattcgtgac atccaccccc tgggcaagtg  
 900

45

agagcatctg caaatgcaa agagaactac ctcatatgca ggatgagcca atggcagctc  
 960  
 caagaaatgt actcggggga caccttacct gtggaagca aatcttttca aaataagcca  
 1020

50

ctgggactcg gtaggtggag cccagctgc tcttctaggg acctatgggg ccttcgtggc  
 1080  
 atctctgtgc tgtgtgctgg ggaggagggt gatgtaatgg tgactctttt ctgatcagca  
 1140

55

ccttgccgt gattccaag gtcccagcca aagcaaaggg ccagttgttt cagtttaaac  
 1200



agacatgtct ttagtctaataaaaattagtt aactgccagt aaagttatit gttagcittg  
1260

5 atgaaagcta tgttggtatc ttccctaata catcaaagta aataaaaaat catttctatg  
1320

taaaaaaaaa aaaaaaaaaa  
1339

10 <210> 30  
<211> 4250  
<212> DNA  
<213> human

15 <400> 30  
gaacacatcg cgtttgcatc ccagaaagta gtcgccgca ctatttcccc caaagagaca  
60

agcacacatg taggaatgac aaaggcttgc gaaggagaga ggcagcccg cgcccgag  
120

20 agatccctc gataatggat tactaaatgg gatacacgct gtaccagttc gtcaggagcc  
180

ccggccgct gtcgctgat gcaccgaaa ggtgaagta gagaaataaa gtctccccgc  
240

25 tgaactacta tgaggtcaga agccttgctg ctatatttca cactgctaca ctttctggg  
300

gctggtttcc cagaagattc tgagccaatc agtatttcgc atggcaacta taaaaacag  
360

30 tatccggtgt ttgtgggcca caagccagga cggaacacca cacagaggca caggctggac  
420

atccagatga ttatgatcat gaacggaacc ctctacattg ctgctaggga ccatatttat  
480

35 actgttgata tagacacatc acacacgga gaaatttatt gtagcaaaaa actgacatgg  
540

aaatctagac aggccgatgt agacacatgc agaataaagg gaaaacataa ggatgagtg  
600

40 cacaacttta ttaaagttct tctaaagaaa aacgatgatg cattgtttgt ctgtggaact  
660

aatgccttca acccttcctg cagaaactat aagatggata cattggaacc attcggggat  
720

45 gaattcagcg gaatggccag atgcccatat gatgccaac atgccaacgl tgcactgttt  
780

gcagatggaa aactatactc agccacagtg actgacttcc ttgccattga cgcagtcatt  
840

50 taccggagtc ttggagaaag ccctaccctg cggaccgtca agcacgattc aaaatggttg  
900

aaagaacat actttgttca agccgtggat tacggagatt atatctactt cttcttcagg  
960

55 gaaatagcag tggagtataa caccatggga aaggtagttt tccaagagt ggctcaggtt

1020  
 tgtaagaatg atatgggagg atctcaaaga gtcctggaga aacagtggac gtcgttcctg  
 1080  
 5 aaggcgcgct tgaactgctc agttcctgga gactctcatt tttatttcaa cattctccag  
 1140  
 gcagttacag atgtgattcg tatcaacggg cgtgatgttg tcctggcaac gttttctaca  
 1200  
 10 ccttataaca gcatccctgg gtctgcagtc tgtgcctatg acatgcttga cattgccagt  
 1260  
 gtttttactg ggagattcaa ggaacagaag tctcctgatt ccacctggac accagttcct  
 1320  
 15 gatgaacgag ttcctaagcc caggccaggt tgctgtgctg gctcatcctc cttagaaaga  
 1380  
 tatgcaacct ccaatgagtt ccctgatgat accctgaact tcatcaagac gcacccgctc  
 1440  
 20 atggatgagg cagtgccctc catcttcaac aggccatggt tcctgagaac aatggtcaga  
 1500  
 taccgcctta ccaaaattgc agtggacaca gctgctgggc catatcagaa tcacactgtg  
 1560  
 25 gtttttctgg gatcagagaa gggaatcatc ttgaagtttt tggccagaat aggaaatagt  
 1620  
 ggttttctaa atgacagcct tttcctggag gagatgagtg tttacaactc tgaaaaatgc  
 1680  
 30 agctatgatg gagtcgaaga caaaaggatc atgggcatgc agctggacag agcaagcagc  
 1740  
 tctctgtatg ttgcgttctc tacctgtgtg ataaaggttc cccltggccg gtgtgaacga  
 1800  
 35 catgggaagt gtaaaaaaac ctgtattgcc tccagagacc catattgtgg atggataaag  
 1860  
 gaagggtggtg cctgcagcca tttatcacc cccagcagac tgacttttga gcaggacata  
 1920  
 40 gagcgtggca atacagatgg tctgggggac tgtcacaatt cctttgtggc actgaatgac  
 1980  
 atttcaactc ctctaccaga taatgaaatg tcttacaaca cagtglatgg gcattccagt  
 2040  
 45 tccctcttgc ccagcacaac cacatcagat tcgacggctc aagaggggta tgagtctagg  
 2100  
 ggaggaatgc tggactggaa gcatctgctt gactcacctg acagcacaga ccctttgggg  
 2160  
 50 gcagtgtctt ccataatca ccaagacaag aaggaggatga ttcgggaaag ttacctcaaa  
 2220  
 ggccacgacc agctggttcc cgtcacctc ttggccattg cagtcatcct ggctttcgtc  
 2280  
 55 atgggggccc tcttctcggg catcacctc tactgcgtct gtgatcatcg gcgcaaagac

2340

gtggctgtgg tgcagcgcaa ggagaaggag ctacccact cgcgccgggg ctccatgagc  
2400

5 agcgtcacca agtcagcgg cctctttggg gacactcaat ccaaagaccc aaagccggag  
2460

gccatcctca cgccactcal gcacaacygc aagctcgcca ctcccggcaa cacggccaag  
2520

10 atgctcatta aagcagacca gcaccacctg gacctgacgg ccctcccac ccagagtca  
2580

acccaacgc tgcagcagaa gcggaagccc agccgcggca gccgcgagtg ggagaggaac  
2640

15 cagaacctca tcaatgcctg cacaaggac atgccccca tgggtcccc tgtgattccc  
2700

acggacctgc ccctgcgggc ctcccacgc cacatccca gcgtggtggt cctgccatc  
2760

20 acgcagcagg gctaccagca tgagtacgtg gaccagcca aaatgagcga ggtggcccag  
2820

atggcgctgg aggaccaggc cgccacactg gagtataaga ccatcaagga acatctcagc  
2880

25 agcaagagtc ccaaccatgg ggtgaacctt gtggagaacc tggacagcct gcccccaaa  
2940

gttccacagc gggaggcctc cctgggtccc ccgggagcct ccctgtctca gaccgtctc  
3000

30 agcaagcggc tggaaatgca ccactcctct tcctacgggg ttgactataa gaggagctac  
3060

cccacgaact cgctcacgag aagccaccag gccaccactc taaaagaaa caactactac  
3120

35 tcctccaatt cctctcacct ctccagaaac cagagctttg gcaggggaga caaccgccg  
3180

cccgcccgcc agagggtgga ctccatccag gtgcacagct ccagccatc tggccaggcc  
3240

40 gtgactgtct cgaggcagcc cagcctcaac gcctacaact cactgacaag gtcggggctg  
3300

aagcgtacgc cctcgctaaa gccggacgta cccccaaac catcctttgc tcccctttcc  
3360

45 acatccatga agcccaatga tgcgtgtaca taatcccagg gggagggggt cagggtgtga  
3420

accagcaggc aaggcgaggc gcccgctcag ctgagcaagg ttctcaactg cctcgagtac  
3480

50 ccaccagacc aagaaggcct gcgagcagc cgaggacgct gggctctcct ctctgggaca  
3540

caggggtact cacgaaaact gggccgcgtg gtttgggtgaa ggtttgcaac ggcggggact  
3600

55 caccttcatt ctcttccttc actttcccc acaccctaca acaggtcgga cccacaaaag

3660  
 acttcagtta tcatcacaaa catgagccaa aagcacatac ctaccccatc cccaccccc  
 3720  
 5 acacacacac acacatgcac acaacacata cacacacacg cacagagggtg aacagaaact  
 3780  
 gaaacatttt gtccacaact tcacgggacg tggccagact gggtttgcgt tccaacctgc  
 3840  
 10 aaaaacacaaa tacattttttt aaaatcaaga aaatttaaaa agacaaaaaa aaaagaattc  
 3900  
 attgataatt ctaactcaga cttaacaat ggcagaagtt tactatgcgc aaalactgtg  
 3960  
 15 aaatgcccg cagtgttaca gctttctgtt gcagcagala aatgccatgt tgggcaacta  
 4020  
 tgtcatagat ttctgctcct cctctctttt aatgaaataa cgtgaccgtt aacgcaagta  
 4080  
 20 actctttatt tattgttcac cctttttttt ctttaaggaaa ggactcttcc aaatatcatc  
 4140  
 ctatgaacag ctcttcagaa agcccatgga aagttaaact atttaacgtg aaatccatta  
 4200  
 25 actggaataa ttgagtttct ttattttttac aataaattca ctgagtaaat  
 4250  
 <210> 31  
 30 <211> 2705  
 <212> DNA  
 <213> human  
 <400> 31  
 ctttagccca acagtcaaaa ataattgatg ctaccctaca aatgtccaaa actctagtat  
 60  
 35 atcatatttc taagttacag caaatattag tcctgctaaa ccaggagct ttggcaaaaa  
 120  
 tgttttttga cagtaaatat gtccttgatt atatattaac tagtcaaaga ggtgtttgta  
 180  
 40 acattattag agcttcttgt ttaggtggg ttaacaccac caatcaagag gtcattctaa  
 240  
 cagaaagcct ggatcagaaa accatcacc taaaaaaca tgccttacat atttaacaca  
 300  
 45 ctctgaaatc cagtcaaaat atgactaaag gcccttgcca tgactgatgt attctcctgg  
 360  
 ccaacgccaa acaaatggga gcctggttac gagtcagcct tcagggactt gtcacatttc  
 420  
 50 tacttggttt ctctcctgtt attgtcataa taaaatgttt tctatgctgt ttagtgcaac  
 480  
 ttaggcccta ttctgtagaa gtctcctcta ctattcaggc cactcaaaca ccccaataa  
 540  
 55

EP 1 355 149 A2

ttgagttcaa aatcgacatc aagatataaa ggaatcagtg actaaatata ttcatatat  
 600  
 5 ggtattttta ttgattattg tgctgtcttg acctagtagg gaggccttgg ctgagggtg  
 660  
 gtcagtttcc tctcttgagc agctgattaa atccacaccc caaccacttc ccttatcagg  
 720  
 10 ttctcacact ctggggccac tatgtaccca ctctaatac cacagggcca gacatcagac  
 780  
 aattaaggac agcgcccatg ccccaaagcc cgccaaaatt atgcaaatta ttcaaaatta  
 840  
 15 ttcaacctag ctaacccac cctttttgct gtacataagc tgccattcc cctccagcc  
 900  
 tgtggtaccc agtcctcagg tgcaacccc tgcgtgttcc tctgtggcag ccttctctca  
 960  
 20 ttcagagctg tttccacag aggtagtga aagaactgga tttcaagtt cactttgcaa  
 1020  
 gagaaaaaga aaactcagta gaagataatg gcaagtccag actggggata tgatgacaaa  
 1080  
 25 aatggctctg aacaatggag caagctgtat ccatttgcca atggaaataa ccaatccct  
 1140  
 gttgatatta aaaccagtga aaccaaacat gacacctctc tgaacctat tagtgtctcc  
 1200  
 30 tacaaccag ccacagcaa agaaattatc aatgtggggc attctttcca tgtaatttt  
 1260  
 gaggacaacg ataaccgatc agtctgaaa ggtggtcctt tctctgacag ctacaggctc  
 1320  
 35 tttcagtttc attttcactg gggcagtaca aatgagcatg gttcagaaca tacagtggat  
 1380  
 ggagtcaaat attctgccga gcttcacgta gctcactgga attctgcaa gtactccagc  
 1440  
 40 cttgctgaag ctgcctcaaa ggctgatggt ttggcagtta ttggtgttt gatgaagggt  
 1500  
 ggtgaggcca acccaaagct gcagaaagta cttgatgcc tccaagcaat taaaaccaag  
 1560  
 45 ggcaaacgag cccattcac aaattttgac cctctactc tccttcctc atccctggat  
 1620  
 ttctggacct accctggctc tctgactcat cctcctctt atgagagtg aacttggatc  
 1680  
 50 atctgtaagg agagcatcag tgtcagctca gagcagctgg cacaattccg cagccttcta  
 1740  
 tcaaatgttg aaggtgataa cgctgtcccc atgcagcaca acaaccgccc aaccaacct  
 1800  
 55 ctgaagggca gaacagtgag agcttcattt tgatgattct gagaagaaac ttgtccttc  
 1860

EP 1 355 149 A2

tcaagaacac agccctgctt ctgacataat ccagttaaaa taataatttt taagaaataa  
1920

5 atttatttca atattagcaa gacagcatgc cttcaaatca atctgtaaaa ctaagaaact  
1980

taaatttttag ttcttactgc ttaattcaaa taataattag taagctagca aatagtaatc  
2040

10 tgtaagcata agcttatctt aaattcaagt ttagtttgag gaattcttla aaattacaac  
2100

taagtgtatt gtatgtctat ttttttcagt ttattlgaac caataaaata attttatctc  
2160

15 tttctttctg ttgtgcattc agttttctaaa accattaagt ttctactcca ttacattca  
2220

aaaatcltaa atactttact tgcaagagta ttttgcttca aatacaacaa cctaagagca  
2280

20 gctggagatg aaatattggg aaattcattt gcttactcct gaagacaaaa atatagctga  
2340

gatgaccact ggatttaata tcgttatgct ggcccaacat tgctaccatt tgtgtgtct  
2400

25 gtgatcaaaa tgattatctt ttatatagga agatgacgct tctggatatt gctttcactt  
2460

cttctcccca cgttagcaag gacaatgctt ctctgccatt attacaacta gttagtllgc  
2520

30 atggagaatc tttactttaa aattggaaga aaagtcacaa gtgaatgggt tataaaaaatg  
2580

ctaaagaagt cattcttgct tagaatcata tagaaacatc atgcaatctt ttagtcagat  
2640

35 gtgcgcttca ccttatgcta tttttatctt taattgacac acaataattg tacatgttta  
2700

tggagtatag tgtggtgttt tctgtttgtt tgtttgtttt ttgagacaag gtctcactct  
2760

40 gccagtcagg gtggagtgcg atggt  
2785

<210> 32  
<211> 9588  
<212> DNA  
45 <213> human

<400> 32  
ccgaccaaca ccaacaccca gctccgacgc agctcctctg cgcccttgcc gccctccgag  
60

50 ccacagcttt cctcccgttc ctgcccccg cccgtcgccg tctccgcgt cgcagcggcc  
120

tcgggagggc ccaggtagcg agcagcgacc tcgcgagcct tccgcactcc cgcccgggtc  
180

55 cccggcgcgc cgctatcct tgccccctc cgctttctcc gcgcgggcc gccctcgctta  
240

tgctcggcg ctgagccgct ctcccgattg cccgccgaca tgagctgcaa cggaggctcc  
 300  
 5 caccgcgga tcaacactct yggccgcatg atccgcgcg agtctggccc ggacctgcgc  
 360  
 tacgaggta ccagcggcg cgggggcacc agcaggatgt actattctcg gcgcggcgtg  
 420  
 10 atcaccgacc agaactcgga cggctactgt caaacggca cgaagtccag gcaccagaac  
 480  
 cagaacacca tccaggagct gctgcagaac tgctccgact gcttgatgag agcagagctc  
 540  
 15 atcgtgcagc ctgaattgaa gtatggagat ggaatacaac tgactcggag tcgagaattg  
 600  
 gatgagtgt ttgccaggc caatgaccaa atggaatcc tcgacagctt gatcagagag  
 660  
 20 atgcggcaga tgggccaggc ctgtgatgt taccagaaaa ggcttcttca gctccaagag  
 720  
 caaatgcgag ccctttataa agccatcagt gtccctcgag tccgcagggc cagctccaag  
 780  
 25 ggtggtggag gctacacttg tcagagtggc tctggctggg atgagttcac caaacatgtc  
 840  
 accagtgaat gtttggggtg gatgaggcag caaagggcgg agatggacat ggtggcctgg  
 900  
 30 ggtgtggacc tqgcctcagt ggagcagcac attaacagcc accggggcat ccacaactcc  
 960  
 atcggcgact atcgtggca gctygacaaa atcaaagccg acctgcgca gaaatctgcg  
 1020  
 35 atctaccagt tggaggagga gtatgaaaac ctgctgaaag cgtcctttga gaggatggat  
 1080  
 cacctgcgac agctgcagaa catcatcag gccacgtcca gggagatcat gtggatcaat  
 1140  
 40 gactgcgagg aggaggagct gctgtacgac tggagcgaca agaacaccaa catcgctcag  
 1200  
 aaacaggagg ccttctccat acgcatgagt caactggaag ttaaagaaaa agagctcaat  
 1260  
 45 aagctgaaac aagaaagtga ccaactgtc ctcaatcagc atccagcttc agacaaaatt  
 1320  
 gaggcctata tggacactct gcagacgag tggagttgga ttcttcagat caccaagtgc  
 1380  
 50 attgatgttc atctgaaaga aaatgctgcc tactttcagt tttttgaga ggcgagctct  
 1440  
 actgaagcat acctgaagg gctccaggac tccatcagga agaagtacc ctgcgacaag  
 1500  
 55 aacatgcccc tgcagcacct gctggaacag atcaaggagc tggagaaaga acgagagaaa  
 1560

atccttgaat acaagcglca ggtgcagaac ttggtaaaca agtctaagaa gattgtacag  
 1620  
 5 ctgaagcctc gtaaccaga ctacagaagc aataaaccca ttattctcag agctctctgt  
 1680  
 gactacaac aagatcagaa aatcgtgcat aagggggatg agtgtatcct gaaggacaac  
 1740  
 10 aacgagcgca gcaagtggta cgtgacgggc ccgggaggcg ttgacatgct tgttccctct  
 1800  
 gtggggctga tcatccctcc tccgaaccca ctggcgtgg acctctcttg caagattgag  
 1860  
 15 cagtactacg aagccatctt ggctctgtgg aaccagctct acatcaacat gaagagcctg  
 1920  
 gtgtctctggc actactgcat gattgacata gagaagatca gggccatgac aatcgccaag  
 1980  
 20 ctgaaaacaa tgcggcagga agattacatg aagacgatag ccgaccttga gttacattac  
 2040  
 caagagttca tcagaaatag ccaaggctca gagatgtttg gagatgatga caagcggaaa  
 2100  
 25 atacagtctc agttcaccga tgcccagaag cattaccaga ccttggctcat tcagctccct  
 2160  
 ggctatcccc agcaccagac agtgaccaca actgaaatca ctcacatgag aacctgcca  
 2220  
 30 gatgtcaacc ataataaagt aattgaaacc aacagagaaa atgacaagca agaaacatgg  
 2280  
 atgctgatgg agctgcagaa gattcgcagg cagatagagc actgcgaggg caggatgact  
 2340  
 35 ctcaaaaacc tccctctagc agaccagggg tcttctcacc acatcacagt gaaaattaac  
 2400  
 gagcttaaga gtgtgcagaa tgattcaca gcaattgctg aggttctcaa ccagcttaaa  
 2460  
 40 gatatgcttg ccaacttcag aggttctgaa aagtactgct atttacagaa tgaagtattt  
 2520  
 ggactatttc agaaactgga aaatatcaat ggtgttacag atggctactt aaatagctta  
 2580  
 45 tgcacagtaa gggcactgct ccaggctatt ctccaaacag aagacatgtt aaaggtttat  
 2640  
 gaagccaggc tcaactgagga ggaaactgtc tgccctggacc tggataaagt ggaagcttac  
 2700  
 50 cgctgtggac tgaagaaaat aaaaaatgac ttgaacttga agaagtcgtt gttggccact  
 2760  
 atgaagacag aactacagaa agcccagcag atccactctc agacttcaca gcagtatcca  
 2820  
 55 ctttatgac tggacttggg caagttcggg gaaaaagtca cacagctgac agaccgctgg  
 2880



caaaggatag ataacagat cgactttaga ttatgggacc lggagaaaca aatcaagcaa  
 2940  
 5     ttgaggaatt atcgtgataa ctatcaggct lctgcaagt ggctctatga tcgtaaacgc  
 3000  
 cgccaggatt ccttagaatc catgaaattt ggagattcca acacagtcac gcggtttttg  
 3060  
 10     aatgagcaga agaacttgca cagtgaata tctggcaaac gagacaaac agaggaagta  
 3120  
 caaaaaattg ctgaactttg cgccaattca attaaggatt atgagctcca gctggcctca  
 3180  
 15     tacacctcag gactggaaac tctgtgaac atacctatca agaggaccat gattcagtc  
 3240  
 ccttctgggg tgattctgca agaggctgca gatgttcacg ctcggtacac tgaactactt  
 3300  
 20     acaagatctg gagactatta caggttctta agtgagatgc tgaagagttt ggaagatctg  
 3360  
 aagctgaaaa ataccaagat cgaagttttg gaagaggagc tcagactggc ccgagatgcc  
 3420  
 25     aactcgaaa actgtaataa gaacaaattc ctggatcaga acctgcagaa ataccaggca  
 3480  
 gagtgtlccc agttcaaagc gaagcttgcg agcctggagg agctgaagag acaggctgag  
 3540  
 30     ctggatggga agtcggctaa gcaaaatcta gacaagtgc acggccaaat aaaagaactc  
 3600  
 aatgagaaga tcacccgact gacttatgag attgaagatg aaaagagaag aagaaaatct  
 3660  
 35     gtggaagaca gatttgacca acagaagaat gactatgacc aactgcagaa agcaaggcaa  
 3720  
 tgtgaaaagg agaactttg ttggcagaaa ttagagtctg agaaagccat caaggagaag  
 3780  
 40     gagtacgaga ttgaaagggt gaggttcta ctgcaggaag aaggcaccgc gaagagagaa  
 3840  
 tatgaaaatg agctggcaaa ggtaagaaac cactataatg aggagatgag taatttaagg  
 3900  
 45     aacaagtatg aaacagagat taacattacg aagaccacca tcaaggagat atccatgcaa  
 3960  
 aaagaggatg attcaaaaa tcttagaaac cagcttgata gactttcaag ggaaaatcga  
 4020  
 50     gatctgaagg atgaaattgt caggctcaat gacagcatct tgcaggccac tgagcagcga  
 4080  
 aggcgagctg aagaaaacgc ccttcagcaa aaggcctgtg gctctgagat aatgcagaag  
 4140  
 55     aagcagcatc tggagataga actgaagcag gtcatgcagc agcgctctga ggacaatgcc  
 4200

cggcacaagc agtccctgga ggaggctgcc aagaccattc aggacaaaaa taaggagatc  
 4260  
 5 gagagactca aagctyagtt tcaggaggag gccaaagcgc gctgggaata tgaaaatgaa  
 4320  
 ctgagtaagg taagaaacaa ttatgatgag gagatcatta gcttaaaaaa tcagtttgag  
 4380  
 10 accgagatca acatcaccaa gaccaccatc caccagctca ccatgcagaa ggaagaggat  
 4440  
 accagtggct accgggctca gatagacaat ctaccccgag aaaacaggag cttatctgaa  
 4500  
 15 gaaataaaga ggctgaagaa cactctaacc cagaccacag agaattctag gaggggtggaa  
 4560  
 gaagacatcc aacagcaaaa ggccactggc tctgaggtgt ctgaggagaa acagcagctg  
 4620  
 20 gaggttgagc tgagacaagt cactcagatg cgaacagagg agagcgttaag atataagcaa  
 4680  
 tctcttgatg atgctgccaa aaccatccag gataaaaaca aggagataga aagggttaaaa  
 4740  
 25 caactgatcg acaaagaaac aaatgaccgg aaatgcctgg aagatgaaaa cgcgagatta  
 4800  
 caaagggctc agtatgacct gcagaaagca aacagtagtg cgacggagac aataaacaaa  
 4860  
 30 ctgaaggctc aggagcaaga actgacacgc ctgaggatcg actatgaaag ggtttcccag  
 4920  
 gagaggactg tgaaggacca ggatatcacg cggttccaga actctctgaa agagctgcag  
 4980  
 35 ctgcagaagc agaaggtgga agaggagctg aatcggctga agaggaccgc gtcagaagac  
 5040  
 tcttgcaaga ggaagaagct ggaggaagag ctggaaggca tgaggaggtc gctgaaggag  
 5100  
 40 caagccatca aaatcaccaa cctgaccag cagctggagc aggcattcat tgttaagaag  
 5160  
 aggagtgagg atgacctccg gcagcagagg gacgtgctgg atggccacct gagggaaaag  
 5220  
 45 cagaggaccc aggaagagct gaggaggctc tcttctgagg tcgaggccct gaggcggcag  
 5280  
 ttactccagg aacaggaaag tgtcaaacaa gctcacttga ggaatgagca tttccagaag  
 5340  
 50 gcgatagaag ataaaagcag aagcttaaat gaaagcaaaa tagaaattga gaggctgcag  
 5400  
 tctctcacag agaacctgac caaggagcac ttgatgttag aagaagaact gcggaacctg  
 5460  
 55 aggctggagt acgatgacct gaggagagga cgaagcgaag cggacagtga taaaaatgca  
 5520

accatcttgg aactaaggag ccagctgcag atcagcaaca accggaccct ggaactgcag  
 5580  
 5 gggctgatta atgatttaca gagagagagg gaaaatttga gacaggaaat tgagaaattc  
 5640  
 caaaagcagg ctttagaggc atctaatagg attcaggaat caaagaatca gtgtactcag  
 5700  
 10 gtggtacagg aaagagagag ccttctggtg aaaatcaaag tcttgagca agacaaggca  
 5760  
 aggcctgcaga ggctggagga tgagctgaat cgtgcaaat caactctaga ggcagaaacc  
 5820  
 15 aggggtgaaac agcgccctgga gtgtgagaaa cagcaaatc agaatgacct gaatcagtgg  
 5880  
 aagactcaat attcccgaag ggaggaggtc attaggaaga tagaatcgga aagagaaaag  
 5940  
 20 agtgagagag agaagaacag tcttaggagt gagatcgaaa gactccaagc agagatcaag  
 6000  
 agaattgaag agaggtgcag gcgtaagctg gaggattcta ccaggagac acagtcacag  
 6060  
 25 ttgaaacag aacgctccc atactcagag gagattgata aactcagaca gcgcccata  
 6120  
 gggctccatc gagagacca gactgagtgt gaggggaccg ttgacacctc caagctggtg  
 6180  
 30 tttgatgggc tgaggaagaa ggtgacagca atgcagctct atgagtgtca gctgatcgac  
 6240  
 aaaacaacct tggacaaact attgaagggg aagaagtcag tggaagaagt tgcttctgaa  
 6300  
 35 atccagccat tccttcgggg tgcaggatct atcgctggag catctgcttc tcctaaggaa  
 6360  
 aaatactctt tggtagaggc caagagaaag aaattaatca gccagaatc cacagtcag  
 6420  
 40 cttctggagg ccagggcagc tacagggtgt ataattgat cccatcgaa tgagaagctg  
 6480  
 actgtcgaca gtgccatagc tcgggacctc attgacttcg atgaccgtca gcagatatat  
 6540  
 45 gcagcagaaa aagctatcac tggttttgat gatccatttt caggcaagac agtatctgtt  
 6600  
 tcagaagcca tcaagaaaaa tttgattgat agagaaaccg gaatgcccct gctggaagcc  
 6660  
 50 cagattgctt cagggggtgt agtagaccct gtgaacagtg tctttttgcc aaaagatgtc  
 6720  
 gccttgcccc gggggctgat tgatagagat ttgtatcgat cctgaatga tccccgagat  
 6780  
 55 agtcagaaaa actttgtgga tccagtcacc aaaaagaagg tcagttacgt gcagctgaag  
 6840

gaacgggtgca gaatcgaacc acatactggt ctgctcttgc tttcagtaca gaagagaagc  
 6900  
 5 atgtccttcc aaggaatcag acaacctgtg accgtcactg agctagtaga ttctggtata  
 6960  
 ttgagaccgt ccactgtcaa tgaactggaa tctggtcaga tttcttatga cgagggtggt  
 7020  
 10 gagagaatta aggacttcct ccagggttca agctgcatag caggcatata caatgagacc  
 7080  
 acaaaacaga agcttggcat ttatgaggcc atgaaaattg gcttagtccg acctggtact  
 7140  
 15 gctctggagl tgctggaagc ccaagcagct actggcttla tagtggatcc tgtagcaac  
 7200  
 ttgaggttac cagtggagga agcctacaag agaggtctgg tgggcattga gttcaaagag  
 7260  
 20 aagctcctgt ctgcagaacg agctgtcact gggataatg atcctgaaac aggaaacatc  
 7320  
 atctctttgt tccaagccat gaataaggaa ctcctcgaaa agggccacg tattcgtta  
 7380  
 25 ttagaagcac agatcgcaac cggggggatc attgacccaa aggagagcca tcgtttacca  
 7440  
 gttgacatag catataagag gggctatttc aatgaggaac tcagtgagat tctctcagat  
 7500  
 30 ccaagtgatg ataccaaagg attttttgac cccaacactg aagaaaatct tacctatctg  
 7560  
 caactaaaag aaagatgcat taaggatgag gaaacagggc tctgtcttct gcctctgaaa  
 7620  
 35 gaaaagaaga aacagggtgca gacatcaca aagaataccc tcaggaagcg tagagtggtc  
 7680  
 atagttgacc cagaaaccaa taaagaaatg tctgttcagg aggcctacaa gaagggccta  
 7740  
 40 attgattatg aaaccttcaa agaactgtgt gagcaggaat gtgaatggga agaaataacc  
 7800  
 atcacgggat cagatggctc caccaggggtg gtcctggtag atagaaagac aggcagtcag  
 7860  
 45 tatgatattc aagatgctat tgacaagggc cttgttgaca ggaagttctt tgatcagtac  
 7920  
 cgatccggca gcctcagcct cactcaattt gctgacatga tctccttgaa aaatggtgtc  
 7980  
 50 ggcaccagca gcagcatggg cagtgggtgc agcagatgatg tttttagcag ctcccgacat  
 8040  
 gaatcagtaa gtaagatttc caccatatcc agcgtcagga atttaaccat aaggagcagc  
 8100  
 55 tctttttcag acaccctgga agaactgagc cccattgcag ccatctttga cacagaaaac  
 8160

ctggagaaaa tctccattac agaaggtata gagcggggca tcqttgacag catcacgggt  
 8220  
 5 cagaggcttc tggaggtca ggcctgcaca ygtggcatca tccaccaac cacgggccag  
 8280  
 aagctgtcac ttcaggacgc agtctcccag ggtgtgattg accaagacat ggccaccagc  
 8340  
 10 gtgaagcctg ctcagaaagc cttcataggc ttcgaggggtg tgaagggaaa gaagaagatg  
 8400  
 tcagcagcag aggcagtga agaaaaatgg ctcccgtatg aggctggcca gcgcttcctg  
 8460  
 15 gagttccagt acctcacggg aggtcttgtt gacccggaag tgcattggag gataagcacc  
 8520  
 gaagaagcca tccggaagg gttcatagat ggccgcgccg cacagaggct gcaagacacc  
 8580  
 20 agcagctatg ccaaaatcct gacctgcccc aaaaccaa ataaaaatc ctataaggat  
 8640  
 gccataaatc gctccatggt agaagatac actgggctgc gccttctgga agccgcctcc  
 8700  
 25 gtgtcgtcca agggcttacc cagcccttac aacatgtctt cggctccggg gtcccgtcc  
 8760  
 ggctcccgt cgggatctcg ctccgatct cgtccgggt cccgcagtgg gtcccggaga  
 8820  
 30 ggaagctttg acgccacagg gaattcttcc tactcttatt cctactcatt tagcagtagt  
 8880  
 tctattgggc actagtagtc agttgggagt ggttgctata ccttgacttc atttatatga  
 8940  
 35 atttccactt tattaaataa tagaaaagaa aatcccgggtg cttgcagtag agtgatagga  
 9000  
 cattctatgc ttacagaaaa tatagccatg attgaaatca aatagtaaag gctgttctgg  
 9060  
 40 ctttttatct tcttagctca tcttaaataa gcagtacact tggatgcagt gcgtctgaag  
 9120  
 tgctaatacag ttgtaacaat agcacaatc gaacttagga tttgtttctt ctcttctgtg  
 9180  
 45 tttcgatttt tgatcaattc ttttaatttg gaagcctata atacagtttt ctattcttgg  
 9240  
 agataaaaaa taaatggatc actgatattt tagtcattct gcttctcacc taaatatttc  
 9300  
 50 catattctgt attaggagaa aattaccctc ccagcaccag ccccccctc aaaccccaa  
 9360  
 cccaaaacca agcatttttg aatgagtctc cttlagtttc agagtgtgga ttgtataacc  
 9420  
 55 catatactct tcgatgtact tgtttggtt ggtattaatt tgactgtgca tgacagcggc  
 9480

aatcttttct ttggtcaaaag ttttctgttt attttgcttg tcatattcga tgtactttaa  
 9540  
 5 ggtgtcttta tgaagtttgc tattctggca ataaactttt agactttt  
 9588  
 <210> 33  
 <211> 366  
 10 <212> DNA  
 <213> human  
 <220>  
 <221> misc\_feature  
 <222> (351)..(351)  
 15 <223> any kind of base  
 <400> 33  
 gaagtgcctat ttatatattat acaaaaatat tacataattc agttaglatl ggtgacataa  
 60  
 20 ttttagttagt atgggtgata taatgggtcat aatttttagc atctaataaa gatcttttta  
 120  
 tgagtcccat ataaaatatg tgaacaaagc aatcttgtca taagatttgt gatgatttag  
 180  
 25 gagaaagtac tttagataa tttttttctg tctctttgtg aactctctca acagtagtgc  
 240  
 tcttttagatt agagccagca ggtcggccat aacagttttc ttcaaatttg ggcaacagtl  
 300  
 30 atacaaatgc ttgaatttca agacaacata ttaaagggtc tatgaactgg naatctaacc  
 360  
 tgggtt  
 366  
 35 <210> 34  
 <211> 1466  
 <212> DNA  
 <213> human  
 40 <400> 34  
 agccccaagc ttaccacctg caccgggaga gctytgtgtc accatgtggg tcccgttgt  
 60  
 45 ctctctcacc ctgtccgtga cgtggattgg tgctgcacc ctcactctgt ctcggttgt  
 120  
 gggaggctgg gagtgcgaga agcattccca accctggcag gtgcttgtgg cctctcgtgg  
 180  
 50 cagggcagtc tgcggcgtg ttctggtgca ccccagtgg gtccctcacag ctgccactg  
 240  
 catcaggaac aaaagcgtga tcttgcgtgg tcggcacagc ctgtttcctc ctgaagacac  
 300  
 55 aggccaggta tttcaggtca gccacagctt cccacacccg ctctacgata tgagcctcct  
 360

EP 1 355 149 A2

gaagaatcga ttccctcaggc cagggtgatga ctccagccac gacctcatgc tgctccgcct  
420

5 gtcagagcct gccgagctca cggatgctgt gaaggctcatg gacctgccca cccaggagcc  
480

agcactgggg accacctgct acgcctcagg ctggggcagc attgaaccag aggagttctt  
540

10 gaccccaaaag aaacttcagt gtgtggacct ccatgttatt tccaatgacg tgtgtgcgca  
600

agttcacctt cagaagggtga ccaagttcat gctgtgtgct ggacgctgga cagggggcaa  
660

15 aagcacctgc tcgggtgatt ctgggggccc acttgctctgt aatggtgtgc ttcaaggat  
720

cacgtcatgg ggcagtgaac catgtgccct gcccgaaagg ccttccctgt acaccaaggt  
780

20 ggtgcattac cggaagtga tcaaggacac catcgtggcc aaccctgag caccctatc  
840

aacccctat ttagtaaac ttggaacctt ggaaatgacc aggccaagac tcaagcctcc  
900

25 ccagttctac tgacctttgt ccttaggtgt gagggtccagg gttgctagga aaagaaatca  
960

gcagacacag gtgtagacca gagtgtttct taaatggtgt aattttgtcc tctctgtgtc  
1020

30 ctgggggaata ctggccatgc ctggagacat atcactcaat ttctctgagg acacagatag  
1080

gatgggggtgt ctgtgttatt tgtgggttac agagatgaaa gaggggtggg atccacactg  
1140

35 agagagtggg gagtgacatg tgctggacac tgtccatgaa gactgagca gaagctggag  
1200

gcacaacgca ccagacactc acagcaagga tggagctgaa aacataaccc actctgtcct  
1260

40 ggaggcactg ggaagcctag agaaggctgt gagccaagga gggagggtct tcctttggca  
1320

tgggatgggg atgaagtaag gagagggact ggacccctg gaagctgatt cactatgggg  
1380

45 ggaggtgtat tgaagtcctc cagacaaccc tcagatttga tgatttccta gtagaactca  
1440

cagaaataaa gagctgttat actgtg  
1466

50 <210> 35  
<211> 187  
<212> DNA  
<213> human

<400> 35  
55 gatctggtgc attccggtcg acactctcgt ttatttggac tglagctcg acctctatga  
60

ataattactt cagcccccta gtgctcccg gccaaagctcc ttggccaaac ttccacctta  
 120  
 5 gcttctgata agtcttgggc caagctaagc agcatctatc aatcatccct tcagctcctg  
 180  
 attgatc  
 187  
 10  
 <210> 36  
 <211> 2913  
 <212> DNA  
 <213> human  
 15 <400> 36  
 actgggtacc gaggactggg tgtgtttaag gcagacagcc aggtgaggat cccagctact  
 60  
 ggggcctgct gtcactcct gggagtaccc gggggtcagg agcctagggg actcttgac  
 120  
 20 ttcacatcca gccatgctaa ttacactttt tggcaaagga aacagctagg agcagtttct  
 180  
 ttcactccta cagccccgtt ttctcagtgt ttagacctcg aattattact gggctagagg  
 240  
 25 gaaggcagcc tctgaagtgt ggcaggagga ggggaagtct gcctgcatct tgggtgtgtct  
 300  
 gtcagatgcc agcactaata acctggcttc tgtgaggcct gtcagtgtct tcaggaatga  
 360  
 30 aaggggaccc ctgagagggtg ctgagtagca gcaggctgtg aatgctctct acccaccacc  
 420  
 ctacactcct cgttaaagat ggtgctacct gccacacagc agacatctgg tcgctgcaca  
 480  
 35 cccgaaagac cccaaggcag tctgccccct gccacgccac acgccagcac ccacctcct  
 540  
 ggccccctgcc tcggcctccc cagaccagct gcacccagcc cccaacacgc accccttctc  
 600  
 40 cagatgtgtg cagggcctca ttttgagag caaagacaga tgtttcagcc acacgcttta  
 660  
 ttaacttcta aaacctgtgc tcaggacact cttcaacagt catgaaaagt ttgatcactt  
 720  
 45 gccacagtca ggacctttgt gtggggctct gatctgatgt tcggctcat catctcccaa  
 780  
 accagcagtc gtttgtaccc caacctctg ctgaggggt cataccccca aatgattttc  
 840  
 50 ctgatttatg tatttccta caaagggtt tctataccta gcctctgcct ccagcatgag  
 900  
 aagggggaat aggtgagacc catttgccag tagcagacgg ggacctggg gagaaaatgg  
 960  
 55 cagagcctgt tggagactcc ctgtctccag ctgaccagcc aatgggattc ctcttccctc



EP 1 355 149 A2

1020  
cactgtctcc cacaaagtag aagaatcctg gtacatttag cccatgagcc tggcacagat  
1080  
5 ccctatctag acatgaggcc ctttagacat gactttggca ttgaccagcc tgttggcaat  
1140  
gggtcgggga ggcagagggg atgctcacac cagtaattct catccctga atgcttggga  
1200  
10 tcacctgggg agagttcaca aaatactggt gcaggggtcc cacctctgat gatgctgagt  
1260  
ggtgggtctg ggtgttgcc caggcatcat gatgtttcag gccccaggt gacttcttag  
1320  
15 gcagcccagc taagccccta gagccttgca atttccccca aatgacctca gagggcccga  
1380  
tttgaggga atgcctaact tcaggggccg taagaatccc ccaggagca tgtgaaatgc  
1440  
20 agataccagg cccaccccc gagatgagct gaggtgggtc aggggtgaag tgcagggatc  
1500  
agtgtttttc acaagctcca tacctccagg aaatggtgtt gtggttgggc ccgtagaaaa  
1560  
25 cattctgaga gtctgttgc ctgtgccttg gtgcacgtgg ggtggaatcc cagtggccct  
1620  
gccttgagga ggatgtgcat taacgtggta ggggagacag agacagctcc acctgcccc  
1680  
30 tgtcccaccg gggacctcca aaaacttcat ggatgttaga gcaagcagcc atgctgcagc  
1740  
agaggatgag gctggcggat ttagtaagag cctctgtgt ttgggtgag tcttttctt  
1800  
35 agttgccctg tcatctggcc tctggataac ccacctctcc tccctcatcc taaaattaca  
1860  
gatggcgaaa gacggccaca ttagtgaga cccctaaggt cctccaacta ggttgggtcc  
1920  
40 acagtggccc ctggtgcatg gaccacacac tctcttccct cctctggctc aggactacgg  
1980  
tctgaaatta gggagatatg aatgtcttcc ttgaaaactt ctcttcccag tcttcccact  
2040  
45 ttgcttgggg gtccttggtc aaggccagct ttggactagg gcttgttgcg actaccagct  
2100  
gtctcatttt gctgtactgc aaactcaggc ttggttccaa gcttatgggg gccctgtcct  
2160  
50 tcccctagta ggglttgttt tggggtcaca tctggtcata cccttcagag agctcttccc  
2220  
cagcctctac atcagggaga gaggtaggta gggaggagca ttcaaggatt agaagaagga  
2280  
55 ctaaagtaca acagccttgg aggaactgcc aggaactaag ggcgagcact ggagaaggca

EP 1 355 149 A2

2340  
 acctgggacc cctgcgctt ctgagcagga agaccaagac cttcaggggc cctaagcact  
 2400  
 5 gaaaacatca ttctcatcc ccaagccctg gcacccccct gttcttctaa aataattctt  
 2460  
 ttctaggtat ttctgattgc aaaattctgg atgggttcat ccaagctgac ctttgctgtt  
 2520  
 10 ttttccttc ccaacaaggc ctcaattttt ggagccacct tagttggtgc ctaggcagag  
 2580  
 gggcagtcag cagtggltat caggatcctg gctctatggg ttgccttctt cctggtctgt  
 2640  
 15 aaagccctcg caggcaggga cttcttagat agctgcttcc ttagggcatg gcattgtgtg  
 2700  
 ggtggttaat gaatggaaga gagggaatga gtgatcaagg gagggaggag ggagtggagt  
 2760  
 20 ggagatttct catcctttcc tgtaattta tgacatcttc ctgcctatga gtccttgact  
 2820  
 ctggagtttt acaaagcagt cacatttcaa ataaaagtct gggaaagcaa cacatcatcg  
 2880  
 25 ccaactttta attttgctaa ataaggatat tag  
 2913  
 <210> 37  
 30 <211> 1466  
 <212> DNA  
 <213> human  
 <400> 37  
 agccccaagc ttaccacctg caccgggaga gctgtgtgtc accatgtggg tcccgttgtt  
 60  
 35 cttcctcacc ctgtcgtga cgtggattgg tgctgcaccc ctcatcctgt ctcggttgtt  
 120  
 gggaggctgg gagtgcgaga agcattccca accctggcag gtgcttgtgg cctctcgtgg  
 180  
 40 cagggcagtc tgcggcggtg ttctgggtga ccccagtggt gtcctcacag ctgcccactg  
 240  
 catcaggaac aaaagcgtga tcttgctggg tcggcacagc ctgtttcatc ctgaagacac  
 300  
 45 aggccaggta ttccaggtca gccacagctt cccacacccg ctctacgata tgagcctcct  
 360  
 gaagaatcga ttctcaggc cagggtgatga ctccagccac gacctcatgc tgctccgctt  
 420  
 50 gtcagagcct gccgagctca cggatgctgt gaaggatcatg gacctgccca cccaggagcc  
 480  
 agcactyggg accacctgct acgcctcagg ctggggcagc attgaaccag aggagtctt  
 540  
 55

EP 1 355 149 A2

gaccccaaaag aaacttcagtc gtgtggaccc ccatgttatt tccaatgacg tgtgtgcgca  
600

5 agttcaccct cagaagggtga ccaagttcat gctgtgtgct ggacgctgga caggggggcaa  
660

aagcacctgc tcgggtgatt ctgggggccc acttgtctgt aatgggtgtc ttcaagggtat  
720

10 cactcatggt ggacgtgaac catgtgccct gcccgaaagg ccttcctgt acaccaaggt  
780

ggtgcattac cggaagtgga tcaaggacac catcgtggcc aaccctgag caccctatc  
840

15 aacccttat tgtagtaaac ttggaacctt ggaaatgacc aggccaagac tcaagcctcc  
900

ccagttctac tgacctttgt ccttaggtgt gaggtccagg gttgctagga aaagaaatca  
960

20 gcagacacag gtgtagacca gagtgtttct taaatggtgt aattttgtcc tctctgtgtc  
1020

ctggggaata ctggccatgc ctggaqacat atcactcaat ttctctgagg acacagatag  
1080

25 gatggggtgt ctgtgttatt tgtgggttac agagatgaaa gaggggtggg atccacactg  
1140

agagagtgga gagtgcacat tgctggacac tgtccatgaa gcactgagca gaagctggag  
1200

30 gcacaacgca ccagacactc acagcaagga tggagctgaa aacataacct actctgtcct  
1260

ggaggcactg ggaagcctag agaaggctgt gagccaagga gggagggtct tcctttggca  
1320

35 tgggatgggg atgaagtaag gagagggaact ggacccctg gaagctgatt cactalgggg  
1380

ggagggtgtat tgaagtcctc cagacaacct tcagatttga tgatttccta gtagaactca  
1440

40 cagaaataaaa gagctgttat actgtg  
1466

<210> 38  
<211> 462  
<212> DNA  
<213> human

45 <220>  
<221> misc feature  
<222> (197)..(197)  
<223> any kind of base

50 <220>  
<221> misc feature  
<222> (116)..(116)  
<223> any kind of base

55

5 <220>  
 <221> misc\_feature  
 <222> (334)..(334)  
 <223> any kind of base  
  
 10 <220>  
 <221> misc\_feature  
 <222> (402)..(402)  
 <223> any kind of base  
  
 15 <220>  
 <221> misc\_feature  
 <222> (429)..(429)  
 <223> any kind of base  
  
 20 <220>  
 <221> misc\_feature  
 <222> (438)..(438)  
 <223> any kind of base  
  
 25 <220>  
 <221> misc\_feature  
 <222> (443)..(443)  
 <223> any kind of base  
  
 30 <220>  
 <221> misc\_feature  
 <222> (459)..(459)  
 <223> any kind of base  
  
 35 <400> 38  
 taagggttta taattatttt tatttttctt ttctttttt tttatggctt ggatgacact  
 60  
 tttttttcag atccaatact agaagttgtt tccatgttca cattttcctt cctggnttaa  
 120  
 aaaaaagagt tgtatttttt ttttttgctt tttttaaat atactttaag ttttagggta  
 180  
 catgtgcaca acgcagnngt tagctacata tgtatacatg tgccatgttg gcgtgctgca  
 240  
 tccagtaact cgtcatttaa cattagglat atctccaaat gctatccttc cccccattgt  
 300  
 atttttcata gcttaaaaat cattgacata ggantaattc caactaaagt acggtattaa  
 360  
 atccctgggg gaataaattt tgtcttaaca agggtaaggt tngtgaaaag gatgggtttg  
 420  
 tcacagggna aaaggganat ccnccattt taaaaccnc ct  
 462  
  
 55 <210> 39  
 <211> 1490  
 <212> DNA  
 <213> human

EP 1 355 149 A2

<400> 39  
ctcgtgcccc ccacggaggg gactgctctc ccccgctgca tcctttctgt gaggtacctt  
60

5 acccacctca gcacctgaga gggtgaaata gaattctaac ctgcacattc gggaagtgtt  
120

tttgagaagt ctcggtcggg aagggaagtc ttccaagtcc gtgcagcact aacgtattgg  
180

10 cacctgcctc ctcttcggcc accccccaga tgaggcagct gtgactgtgt caagggaagc  
240

cacgactctg accatagtct tctctcagct tccactgccg tctccacagg aaaccagaa  
300

15 gttctgtgaa caagtccatg ctgccatcaa ggcatttatt gcagtgtact attgcttcc  
360

aaaggatcag gggatcaccg tgagaaagct ggtacggggc gccaccctyg acatcgtgga  
420

20 tggcatggct cagctcatgg aagtactttc cgtcactcca actcagagcc ctgagaacaa  
480

tgaccttatt tcctacaaca gtgtctgggt tgcgtgccag cagatgcctc agataccaag  
540

25 agataacaaa gctgcagctc ttttgatgct gaccaagaat gtggattttg tgaaggatgc  
600

acatgaagaa atggagcagg ctgtggaaga atgtgacctt tactctggcc tcttgaatga  
660

30 tactgaggag aacaactctg acaaccacaa tcatgaggat gatgtgttgg gggttcccag  
720

caatcaggac ttgtattggt cagaggacga tcaagagctc ataatcccat gccttgcgct  
780

35 ggtgagagca tccaaagcct gcctgaagaa aattcggatg ttagtggcag agaatgggaa  
840

gaaggatcag gtggcacagc tggatgacat tgtggatatt tctgatgaaa tcagccctag  
900

40 tgtggatgat ttggctctga gcatatatcc acctatgtgt cacctgaccg tgcgaatcaa  
960

ttctgcgaaa cttgtatctg ttttaagaa ggcacttgaa attacaaaag caagtcattg  
1020

45 gacccctcag ccagaagata gttggatccc ttacttatt aatgccattg atcattgcat  
1080

gaatagaatc aaggagctca ctccagagtga acttgaatta tgacttttca ggctcatltg  
1140

50 tactctcttc ccctctcctc gtcattgtca ggctctgata cctgccttta aaatggagct  
1200

agaatgcttg ctggattgaa agggagtgc tctctatatt tagcaagaga cactattacc  
1260

55 aaagattgll ggtaggcca gattgacacc tatttataaa ccatatgcgt atatttttct

EP 1 355 149 A2

1320

gtgctatata tgaaaaataa ttgcatgatt tctcattcct qagtcatttc tcagagattc  
1380

ctaggaaagc tgccttattc tclttttgca gtaaagtatg ttgttttcat tgtaaagaty  
1440

ttgatgtgtc caataaaatg ctaacttgcc agtgattaaa aaaaaaaaaa  
1490

<210> .40  
<211> 1677  
<212> DNA  
<213> human

<400> 40  
cttgacccta tttatagtgg ctctaaaggt ggtgttatta tgttttctag agcacttcga  
60

ttatacaaac gtcaaggaat ccgagttaat gtgctttgcc ctgagtttgt tgaaacagac  
120

atgggcacaa tgatcgggtcc caaatccctt agtatgatgg ggggctttgt acctatggaa  
180

atgggtggtga aaggtgcttt tgagctcatc actgatgaga ataaagccgg cgattgccta  
240

tggattacta atcggcgagg tcttgagtac tggcccaccc catcagaaga agcaaagtac  
300

ttgctgcgtt ctacacgttc caggagaaga actgaataca aagctccacc aattaaacta  
360

cctgagagtt ttgagaaaat agttgttcag accttgactc acaactttcg gaatgtacc  
420

agtgtagtaa gagcaccact gagattacct atcaaaccaa actatgttct tgtgaagata  
480

atctatgctg gtgtaaatgc tagtgatgta aattttagct cagggtcgcta ttttggtggc  
540

aataacagtg aactgcac cgcgttccg tttgatgcag gatttgaggc tgtgggagta  
600

attgcagcag ttggggattc tgttactgac ttgaaagttg gcatgccttg tgcgttcag  
660

acttttgag gctatgctga atttacaatg attccttcga aatacgccct tccaatgcct  
720

agaccagaac cggaaggtgt tgccatgctt acatcaggat taacagcttc aattgctcta  
780

gaaaaggcag gacagatgga atctggaaaa gtggtccttg ttactgctgc ggcaggagga  
840

actggtcagt ttgctgttca gcttgcaaaa ttagctggta ataccgtggt tgccacttgt  
900

ggaggtgggg caaaggccaa gcttctgaaa gaattgggag tcgacagagt catagactat  
960

EP 1 355 149 A2

cacagtgaag atataaaaac ggttctaagg aaagagttcc cgaaaggat tgatatcatc  
 1020  
 5      tacgaatctg ttggtgggga catgttaaag ttgtgcttgg atgctttggc agtccatgga  
 1080  
 cgactcattg tcattggcat gatttctcag tatcaaggag aaaatggttg gacgccatca  
 1140  
 10      aaatatcctg gactatgtga gaagctcttg tcaaagagtc aaactgtggc tggctttttc  
 1200  
 ctggtgcaat atagtcacat gtaccaagaa caccttaaca agttatttga cctttactct  
 1260  
 15      tccggaaaaac taaagggttg tgtggatcca aagagattta taggccttca ttctgttgct  
 1320  
 gatgctgttg agtatctcca ttcaggcaaa agcgttggga aggtggttgt ctgcgtggac  
 1380  
 20      ccgaccttcg gtcacaaagt agccaaatta tgaatgaaca cgggtgcaaa tacagaaaga  
 1440  
 agtgaagttt tcaattctta gtctagagat tgttctcgaa tgttactgaa aatagctgct  
 1500  
 25      agaccagtgc tggaaatattt attctcaatg ctttttcaat ttgggattac ttgaaagaat  
 1560  
 aatccattta tgtataccat gtttatgttt acactatata acaactatga gcagaagaaa  
 1620  
 30      gcgagatata taaaaataa attataatcc ttccatttta aaaaaaaaaa aaaaaaa  
 1677  
 <210> 41  
 <211> 1330  
 <212> DNA  
 35      <213> human  
 <400> 41  
 atggcgcagt gggacagctt cactgatcaa caggaggaca ctgatagctg ttcagaatct  
 60  
 40      gtgaagtttg atgctcgctc caatacagct ttgcttcccc caaatcctaa aaatggccct  
 120  
 ccacttcaag aaaagctgaa atccttcaaa gctgcactga ttgcccttta tctccttggtg  
 180  
 45      tttgctgttc tcattcctat catcgcaata atggcagctc aactcctgaa gtgggaaatg  
 240  
 aagaattgca cagttggttc aattaatgca aacagtgtat cctccagtct cctgggaaga  
 300  
 50      ggaaatgaca gtgaagatga agtgagattt cgagaagttg ttatggaaca cattagcaag  
 360  
 atggagaaaa gaatccaata latttcagat actgaagaaa atctcgtaga ttcagagcat  
 420  
 55      tttcaaaatt tcagtgtgac aactgatcaa cgatttgctg atgttcttct ccaactaagt  
 480

accttggttc ccacagtcca gggacatggg aatgccgtag atgaaatcac caggtcctta  
540

5     ataagtctga ataccacgct gcttgatttg cacctctatg tagaaacact gaatgtcaaa  
600

ttccaggaga atacacttaa agggcaagag gaaatcagca aattaaagga gcgtgtgcac  
660

10    aatgcatcag cagaaattat gtctatgaaa gaagaacaag tgcatttggg acaggaaata  
720

aaaagagaag tgaaagtcct gaataacatc actaatgac tcaggctgaa agattgggaa  
780

15    cattctcaga cggttgagaaa tatcacttta attcaaggct ctccctggacc cccaggagaa  
840

aaaggagata gaggtccaac cggagaaagt ggtccaccag gcgttccagg tccagtaggt  
900

20    cctccaggtc ttaaggggtga tcgaggatct attggctttc cgggaagtcg aggatatcca  
960

ggacaatcag ggaagactgg gaggacagga taccctggac caaaaggcca aaagggaqaa  
1020

25    aaaggcagtg gaagcaccct gactccttct gcgactgtcc gactggtttg tggccgtggc  
1080

cctcatgagg gtagagtggg gataatgcac aatggacagt ggggcacagt ttgtgatgat  
1140

30    cactgggaac tgcgtgccgg gcaggttgtc tgcaggagct tgggataccg aggtgttaag  
1200

agtgtgcaca agaaagctta ttttggacaa ggtactggtc ccatttggct gaatgaagta  
1260

35    ccctgttttg gcatggagtc atccattgaa gaggcagaaa tcagacagtg gggcgtgaga  
1320

gtctgttcac  
1330

40    <210> 42  
       <211> 431  
       <212> DNA  
       <213> human

45    <220>  
       <221> misc feature  
       <222> (97)..(97)  
       <223> any kind of base

50    <220>  
       <221> misc feature  
       <222> (347)..(347)  
       <223> any kind of base

55    <220>  
       <221> misc\_feature



<222> (349)..(349)  
 <223> any kind of base

5        <220>  
          <221> misc\_feature  
          <222> (361)..(361)  
          <223> any kind of base

10       <220>  
          <221> misc\_feature  
          <222> (362)..(362)  
          <223> any kind of base

15       <220>  
          <221> misc\_feature  
          <222> (363)..(363)  
          <223> any kind of base

20       <220>  
          <221> misc\_feature  
          <222> (401)..(401)  
          <223> any kind of base

25       <220>  
          <221> misc\_feature  
          <222> (428)..(428)  
          <223> any kind of base

30       <400> 42  
          ctttttatat ttattttcat cgctacacaa acatttttta ggagtttgat tctacctcca  
          60  
          ttttggttag atatacaaac tctaccccat gagggantgt atggtgtatt tctagattta  
          120

35       gcaacaattt tcttgaaaaa tgtacaatac tatagaaaaa tgaagatagt aaataccagg  
          180  
          tataagttaa taacagtgtt tcttttgttc agtaataatg aactgtgtac tagcactgaa  
          240

40       ctttaggcc ccttatitgc gtattttctg tttgtatatt tttaaataga ggaattgtga  
          300  
          ttataatatt attattttgg aatatoctaa atcataaatt caaaacntna tttagttttt  
          360

45       nnnntttttt tttaagatgg agtcccgtt tgtcccaggc nggagtgcag tggcatgatc  
          420  
          tcagctcnct g  
          431

50       <210> 43  
          <211> 669  
          <212> DNA  
          <213> human

55       <220>

EP 1 355 149 A2

<221> misc\_feature  
 <222> (641)..(641)  
 <223> any kind of base

5

<400> 43  
 ttcttttgga aaaccaaaca tgctttatctt catttttttc acaattttatt taaacatctc  
 60

10

acatatataa aataggtaca atttaatttt tctgcttgcc caagaaacaa agcttctgtg  
 120

gaaccatgga agaagatgaa aatgagactg gcaaagaaca aatgctgaat ctgaagaaga  
 180

15

ggacaacttt gggcaaataa tctgcatact ttttaattggg aataagatgg aaaatatgaa  
 240

tgctaaatca aattttttta aaaaatacacc acacgataca actcaatata ggagtatttc  
 300

20

ttctcaaatt cttctagcac catcaacatt cttcaagtat ctgaaatact attaattagc  
 360

accttttgat tatgaacaaa acaaaacaag gacctcagtt catclctgtc taggtcagca  
 420

25

cctaacaatg tggatcacac tcatgggaaa gtgttttgag gtagttttaa cctttggaag  
 480

tttgggtttt aaacttcctt ctgtggaaga tattcaaaag ccacaagtgg tgcaaatgtt  
 540

30

tatgggtttt atttttcaat ttttattttg gttttcttac aaagggtgac atttttcata  
 600

acaggtgtaa gagtggtgaa aaaaaattt caatttttgg ngggaacggg ggaaggagtt  
 660

35

aatgaaact  
 669

40

<210> 44  
 <211> 287  
 <212> DNA  
 <213> human

<400> 44  
 gccggagagt ctacaatggt acccagcatg ctgttggcat tgttgtaaac aaacaagtta  
 60

45

agggcaagat tcttgccaag agaattaatg tgcgtattga gcacattaag cactctaaga  
 120

gccgagatag cttcctgaaa cgtgtgaagg aaaatgatca gaaaaagaaa gaagccaaag  
 180

50

agaaaggtag ctgggttcaa ctaaagcgcc acgctgctcc acccagagaa gcacactttg  
 240

tgagaaccaa tgggaaggag cctgagctgc tggaacctat tccctat  
 287

55

<210> 45

EP 1 355 149 A2

<211> 383  
 <212> DNA  
 <213> human

5

<220>  
 <221> misc feature  
 <222> (147)..(147)  
 <223> any kind of base

10

<220>  
 <221> misc feature  
 <222> (309)..(309)  
 <223> any kind of base

15

<220>  
 <221> misc feature  
 <222> (349)..(349)  
 <223> any kind of base

20

<220>  
 <221> misc feature  
 <222> (365)..(365)  
 <223> any kind of base

25

<220>  
 <221> misc feature  
 <222> (372)..(372)  
 <223> any kind of base

30

<220>  
 <221> misc feature  
 <222> (380)..(380)  
 <223> any kind of base

35

<400> 45  
 ggaacggaaa aggagaattc aagtgtgacc ctcatgaggc aacgtgttat gatgatggga  
 60  
 agacatacca cgtaggagaa cagtggcaga aggaatatct cggtgccatt tgctcctgca  
 120

40

catgctttgg aggccagcgg ggctcgnctt gtgacaactg ccgcagacct ggggggtgaa  
 180  
 ccagtcctcg aaggcactac tggccagtcc tacaaccagt attcttcaga gataccattc  
 240

45

agagaacaaa cactaatggt taatttgccc aatttgagtg ctcatgcct tttaggatgt  
 300  
 tacaggctng acagagaagg ttttcccgag gagttaaata atctttttnc catttccga  
 360

50

ggggncaagg cntgttttn ttt  
 383

55

<210> 46  
 <211> 523  
 <212> DNA

<213> human  
 <400> 46  
 5 cagaggggca gggcggacgg ctaggagttc aagaaacatc ctggtctgag ggaaaggctg  
 60  
 cagctgcacc gccatgaata agcttttcag cttctggaag aggaagaatg agacccgcag  
 120  
 10 ccagggtac aaccttcgag aaaaggattt aaagaaactt cacagagctg cttcagtcgg  
 180  
 ggatttgaag aagctgaagg aataccttca gatcaagaaa tatgatgtaa atatgcagga  
 240  
 15 ctatgaatac agaacacctt tgcacctagc ctgtgctaata ggacatacag atgttgtact  
 300  
 ttctctaatt gagcaacaat gcaagataaa tgtccgggat agtgaaaaca aatccccatt  
 360  
 20 gattaaggca gtacagtgtc aaaatgagga ttgtgtact attctgctaa acttttgtgc  
 420  
 agaccagat ctgagggata ttcgttataa tactgttctt cactatgctg tttgtgtca  
 480  
 25 aagtttgtca ttagttgaaa aactgcttga atacgaagct gat  
 523  
 <210> 47  
 <211> 390  
 <212> DNA  
 30 <213> human  
 <400> 47  
 tccaaggtca tggcaaaaca tctgaagttc atcgccagga ctgtgatggt acaggaaggg  
 60  
 35 aacgtggaaa gcgcatacag gaccttaaac agaatacctca ctatggatgg gctcattgag  
 120  
 gacattaagc atcggcggta ttatgagaag ccatgccgcc gcgacagagg gaaagctatg  
 180  
 40 aaaggtgccg gcggatctac aacatggaaa tggctcgcaa gatcaacttc ttgatgcgaa  
 240  
 agaatcgggc agatccgtgg cagggtctgt gaggcctgtg ggtgggacac cagtgcgaaa  
 300  
 45 ccctcatcca gtttctctc catctctttt ctttgtacaa tccatttcc tattaccatt  
 360  
 ctctgcaata aactcaaata acatgtctgc  
 390  
 50 <210> 48  
 <211> 669  
 <212> DNA  
 <213> human  
 55 <220>  
 <221> misc\_feature

<222> (641)..(641)  
 <223> any kind of base

5 <400> 48  
 ttcttttggg aaaccaaaca tgctttattt cttttttttc acaatttatt taaacatctc  
 60  
 acatatacaa aataggtaca atttaatttt tctgcttgcc caagaacaa agcttctgtg  
 120  
 10 gaaccatgga agaagatgaa aatgagactg gcaaagaaca aatgctgaat ctgaagaaga  
 180  
 ggacaacttt gggcaaataa tctgcatact ttttaattggg aataagatgg aaaatatgaa  
 240  
 15 tgctaaatca aattttttta aaaatacacc acacgatata actcaatata ggagtatttc  
 300  
 ttctcaaatt cttctagcac catcaacatt cttcaagtat ctgaaatact attaattagc  
 360  
 20 acctttgtat tatgaacaaa acaaaacaag gacctcagtt catctctgtc taggtcagca  
 420  
 cctaacaatg tggatcacac tcatgggaaa gtgttttgag gtagtttaaa cctttggaag  
 480  
 25 tttgggtttt aaacttcctt ctgtggaaga tattcaaaag ccacaagtgg tgcaaatgtt  
 540  
 tatgggtttt atttttcaat ttttattttg gttttcttac aaaggttgac atttttcata  
 600  
 30 acaggtgtaa gagtgttgaa aaaaaaattt caatttttgg ngggaacggg ggaaggagtt  
 660  
 aatgaaact  
 669

35 <210> 49  
 <211> 431  
 <212> DNA  
 <213> human

40 <220>  
 <221> misc\_feature  
 <222> (97)..(97)  
 <223> any kind of base

45 <220>  
 <221> misc\_feature  
 <222> (347)..(347)  
 <223> any kind of base

50 <220>  
 <221> misc\_feature  
 <222> (349)..(349)  
 <223> any kind of base

55 <220>

5 <221> misc\_feature  
 <222> (361)..(361)  
 <223> any kind of base  
  
 <220>  
 <221> misc\_feature  
 <222> (362)..(362)  
 <223> any kind of base  
 10  
  
 <220>  
 <221> misc\_feature  
 <222> (363)..(363)  
 <223> any kind of base  
 15  
  
 <220>  
 <221> misc\_feature  
 <222> (401)..(401)  
 <223> any kind of base  
 20  
  
 <220>  
 <221> misc\_feature  
 <222> (428)..(428)  
 <223> any kind of base  
 25  
  
 <400> 49  
 ctttttatat ttattttcat cgctacacaa acatttttta ggagtttgat tctacctcca  
 30 60  
 ttttggttag atatacaaac tctaccccat gagggantgt atgggtgtatt tctagattta  
 120  
 gcaacaattt tcttgaaaaa tgtacaatac tataaaaaa tgaagatagt aaataccagg  
 35 180  
 tataagttaa taacagtgtt tcttttgttc agtaataatg aactgtgtac tagcactgaa  
 240  
 ctttaggccc tcctatttgc gtattttctg tttgtatatt tttaaataga ggaattgtga  
 40 300  
 ttataatatt attattttgg aatatacctaa atcataaatt caaaacntna tttagttttt  
 360  
 nnntttttt tttaagatgg agtcccgctt tgtcccaggc nggagtgcag tggcatgatc  
 45 420  
 tcagctcnct g  
 431  
 50

# Claims

- 55
1. A method of assessing colorectal cancer status comprising identifying differential modulation of each gene (relative to the expression of the same genes in a normal population) in a combination of genes selected from the group consisting of Seq. ID. No. 42-49.

2. The method of claim 1 wherein there is at least a 2 fold difference in the expression of the modulated genes.
3. The method of claim 1 wherein the p-value indicating differential modulation is less than .05.
- 5 4. The method of claim 1 further comprising employing a colorectal diagnostic that is not genetically based.
5. The method of claim wherein the cancer marker that is not genetically based is selected from the group consisting of carcinoembryonic antigen, CA19-9, CA 125, CK-BB, and Guanylyl Cyclase C.
- 10 6. A diagnostic portfolio comprising isolated nucleic acid sequences, their complements, or portions thereof of a combination of genes selected from the group consisting of Seq. ID. No. 42-49.
7. The diagnostic portfolio of claim 6 in a matrix suitable for identifying the differential expression of the genes contained therein.
- 15 8. The diagnostic portfolio of claim 7 wherein said matrix is employed in a microarray.
9. The diagnostic portfolio of claim 8 wherein said microarray is a cDNA microarray.
- 20 10. The diagnostic portfolio of claim 8 wherein said microarray is an oligonucleotide microarray.
11. A diagnostic portfolio comprising isolated nucleic acid sequences, their complements, or portions thereof of a combination of genes selected from the group consisting of Seq. ID. No. 42-49.
- 25 12. A kit for diagnosing colorectal cancer comprising isolated nucleic acid sequences, their compliments, or portions thereof of a combination of genes selected from the group consisting of Seq. ID. No. 42-49.
13. The kit of claim 12 further comprising reagents for conducting a microarray analysis.
- 30 14. The kit of claim 12 further comprising a medium through which said nucleic acid sequences, their complements, or portions thereof are assayed.
15. A method of assessing response to treatment for colorectal cancer comprising identifying differential modulation of each gene (relative to the expression of the same genes in a normal population) in a combination of genes selected from the group consisting of Seq. ID. No. 42-49.
- 35 16. The method of claim 15 wherein the assessment of the response to therapy includes a determination of whether the patient is improving, not improving, relapsing, likely to improve, or likely to relapse.
- 40 17. Articles for assessing colorectal cancer statuts comprising isolated nucleic acid sequences, their complements, or portions thereof of a combination of genes selected from the group consisting of Seq. ID. No. 42-49.
18. Articles for assessing colorectal cancer status comprising representations of isolated nucleic acid sequences, their complements, or portions thereof of a combination of genes selected from the group consisting of Seq. ID. No. 42-49.
- 45
- 50
- 55